The influence of written input on encoding difficult phonological contrasts for L1 Arabic learners of L2 English

Louise Shepperd

PhD

University of York
Education

June 2023

Abstract

Research investigating the relationship between orthography and second language (L2) phonology has grown considerably over the past 20 years, in both size and complexity (Hayes-Harb & Barrios, 2021). However, few studies explore the influence of L2 written input across writing systems and varying proficiency levels (Hao & Yang, 2021; Mok et al., 2018; Showalter, 2020). Further, participants are rarely invited to reflect on their beliefs and strategies related to written input. The present study draws together behavioural and reflective insights to investigate the influence of written input on the lexical encoding of confusable phonological contrasts by adult first language (L1) Arabic-speaking learners of L2 English.

An L1 Arabic group (n=114), with varying English proficiency, and a control L1 English group (n=117), with no Arabic experience, completed an online word learning study. Word learning involved 12 English pseudowords presented auditorily, accompanied by an image and written input. Words were minimal pairs differing by either /m-n/ (easy) or /f-v/ (difficult) contrasts. Minimal pairs from each contrast were taught with Arabic spelling, English spelling, or audio only. Lexical encoding was then tested in an audio-visual matching task, followed by a reflective post-test questionnaire.

Mixed-effects modelling of L1 Arabic responses revealed an inhibitory effect of any written input on accuracy matching /f-v/ words, while written input facilitated encoding of /m-n/ words. In contrast to the behavioural findings, participant reflections on the influence of written input were overwhelmingly positive. Participants valued written input for a variety of reasons and reported distinct influences of Arabic and English script input. Further, a range of strategies were reported, drawing on phonological and orthographic knowledge; however, these strategies did not improve accuracy. Both quantitative and qualitative analysis converged on the pivotal role of individual differences, such as proficiency, in modulating the relation between written input and phonological learning.

Contents

Abstract **2**
Contents **3**
List of tables **8**
List of figures **11**
Acknowledgements **15**
Author's declaration **16**

Chapter 1: Introduction **17**
1.1 Motivation and contribution 18
1.2 Thesis outline 20

Chapter 2: Literature review **21**
2.1 The role of input in second language acquisition 21
2.1.1 Input-related factors 24
2.1.2 Processing, comprehending, and learning 25
2.1.3 The role of consciousness in relation to input 29
2.1.3.1 Implicit and explicit distinctions 29
2.1.3.2 Awareness, attention, and noticing 32
2.1.3.3 Language learning strategies 34
2.1.4 Summary of the role of input in SLA 36
2.2 The role of literacy and orthography in phonological development 37
2.2.1 L1 phonological awareness and literacy development 39
2.2.2 Models of L2 phonological learning 47
2.2.3 Orthographic input and L2 phonological acquisition 53
2.2.3.1 Orthographic influence on nonnative contrasts in novel words 56
2.2.3.2 Orthographic influence across writing systems 62
2.2.4 Summary of orthographic influence on phonological development 68
2.3 Arabic and English language and literacy 69
2.3.1 Arabic phonology and orthography 69
2.3.2 English phonology and orthography 76
2.4 Rationale and research questions 79

Chapter 3: Methodology **82**
3.1 Approach 82
3.1.1 Statement of positionality 84
3.2 Pilot 86
3.2.1 Pilot study 1 87

3.2.2 Pilot study 2 88
3.3 Participants 88
3.3.1 Recruitment 88
3.3.2 Background information and exclusions 90
3.4 Stimuli design 94
3.4.1 Auditory and written stimuli norming 97
3.4.2 Images and norming 98
3.5 Procedure 100
3.6 Research Instruments 101
3.6.1 Audio check 101
3.6.2 English check 102
3.6.3 Background questionnaire 103
3.6.4 Word learning phase 104
3.6.5 Audio-visual matching task 105
3.6.6 Distraction check 106
3.6.7 Post-test questionnaire 107
3.6.8 Development of the POLLS inventory 108
3.7 Ethical considerations 114
3.8 Internet-based factors 115
3.8.1 Stimuli presentation and response times 116
3.8.2 Recruitment and retention 119
3.8.3 The 'who', 'where' and 'how' of participation 120
3.9 Analysis overview 123

Chapter 4: Influence of written input on lexical encoding of contrasts 124
4.1 Introduction 124
4.2 Research context 125
4.3 Methodology 127
4.4 Word learning results 130
4.4.1 GLMMs with word learning accuracy 131
4.4.2 GLMMs with word learning response times 133
4.4.3 Word learning summary 135
4.5 Word matching results 136
4.5.1 Descriptive overview 136
4.5.2 GLMMs with L1 English-speaker accuracy 142
4.5.3 GLMMs of L1 Arabic-speaker accuracy 148
4.5.4 Word matching summary 154
4.6 Concluding remarks 156

Chapter 5: Participant perspectives on written input..............159
5.1 Introduction..............159
5.2 Methodology..............162
5.3 Analytical Procedure..............166
5.4 Participant perceptions of OI influence during the task..............172
5.4.1 Frequencies of perceived OI influence..............173
5.4.2 QCA of perceived OI influence..............178
5.4.2.1 Written input helps..............187
5.4.2.2 Written input does not help..............193
5.4.3 Summary of perceived OI influence..............195
5.5 Participant beliefs about the importance of written input..............198
5.5.1 Frequencies of perceived OI importance..............198
5.5.2 QCA of perceived OI importance..............199
5.5.2.1 Written input is important..............208
5.5.2.2 Written input is not important..............216
5.5.3 Summary of perceived OI importance..............220
5.6 Concluding remarks..............223

Chapter 6: Participant strategies and written input..............228
6.1 Introduction..............228
6.2 Research context..............229
6.3 Methodology..............233
6.4. Qualitative Analysis..............237
6.4.1 Analytical approach..............237
6.4.2 QCA of language learning strategies..............239
6.4.2.1 Cognitive strategies..............250
6.4.2.2 Metacognitive strategies..............256
6.4.3 QCA strategy summary..............259
6.5 Quantitative Analysis..............262
6.5.1 Descriptive overview of strategy frequencies..............262
6.5.2 PCA with strategy inventory items..............271
6.5.2.1 PCA with L1 Arabic data..............271
6.5.2.2 PCA with L1 English data..............277
6.5.2.3 PCA summary..............283
6.5.3 Relationships between strategies, proficiency, and lexical encoding..............284
6.5.3.1 Correlations analysis of individual differences and accuracy..............285
6.5.3.2 Multiple regression analysis of accuracy and strategies..............287

6.5.3.3 Linear regressions of strategies by proficiency... 288
6.6 Concluding remarks...290

Chapter 7: Beyond proficiency...293
7.1 Introduction...293
7. 2 Research context... 294
7.3 Cluster Analysis of proficiency groupings... 296
7.3.1 Analytical approach...297
7.3.2 Results...299
7.3.3 Summary... 302
7.4 Learner profiles by proficiency group...302
7.4.1 Lower proficiency learner profiles...304
7.4.2 Higher proficiency learner profiles... 311
7.4.3 Summary... 319
7.5 Concluding remarks...321

Chapter 8: Discussion...323
8.1 Introduction...323
8.2 Key findings...324
8.2.1 Influence of written input on lexical encoding of contrasts...326
8.2.2 Participant perspectives on written input... 327
8.2.3 Participant strategies and written input... 328
8.2.4 Beyond proficiency... 329
8.2.5 Summary... 330
8.3 Theoretical contributions...331
8.3.1 Building on models of L2 speech learning... 332
8.3.2 Alternative perspectives on orthographic influence...334
8.3.3 Predicting orthographic influence across writing systems...337
8.4 Pedagogical applications... 342
8.5 Methodological considerations...347
8.6 Summary of future directions...353

Chapter 9: Conclusion... 355

Appendix I: Arabic information and consent...358
Appendix II: English information and consent...363
Appendix III: English and Arabic debriefing information...366
Appendix IV: English and Arabic plain language summaries...369
Appendix V: Data pre-processing Rmd output...371

Appendix VI: Background questionnaire **387**
Appendix VII: Post-test questionnaire **391**
Appendix VIII: Matching task data exploration Rmd output **394**
Appendix IX: Word learning Rmd output **412**
Appendix X: Word matching GLMM Rmd output **427**
Appendix XI: Coder guidelines for Q6 .... **445**
Appendix XII: Coder guidelines for Q8 .... **456**
Appendix XIII: Intercoder reliability calculations Q6 and Q8 .... **468**
Appendix XIV: Coded questionnaire responses for Q6 .... **470**
Appendix XV: Coded questionnaire responses for Q7 and Q8 .... **485**
Appendix XVI: Coder guidelines for Q13 .... **501**
Appendix XVII: Intercoder reliability calculations Q13 .... **512**
Appendix XVIII: Coded questionnaire responses for Q13 .... **514**
Appendix XIX: LLS data pre-processing Rmd output **531**
Appendix XX: LLS PCA Rmd output **533**
Appendix XXI: LLS correlational and regression Rmd output **556**
Appendix XXII: Cluster analysis Rmd output **567**
Appendix XXIII: Teaching resource for assessing orthographic facilitation/interference **574**
Appendix XXIV: Modified POLLS Inventory **575**

Glossary **579**
References **581**

List of tables

Table 2.1: Unified Competition Model risk and support factors for second language learning (MacWhinney, 2012, p. 34)..27

Table 2.2: Overview of MSA and British English consonant inventories..................................73

Table 2.3: Overview of Arabic graphemes, letter forms and phoneme correspondence........75

Table 2.4: Example grapheme-phoneme correspondences and letter forms for English consonants, including examples with varying word positioning...77

Table 3.1: Overview of participant background information...91

Table 3.2: Qualitative analysis of master's dissertation questionnaire responses when asked: "What strategies did you use to try to learn the words?"..109

Table 3.3: Comparison of strategy categories and statements for the POLLS inventory......111

Table 3.4: Overview of software and hardware tested for timing by large-scale studies.....117

Table 4.1: Mean accuracy and response times from word learning phase, by L1 groups....130

Table 4.2: GLMM model summary for word learning accuracy...132

Table 4.3: GLMM model summary for word learning response times..................................135

Table 4.4: Summary of mean % correct responses in match and mismatch trials, *d′* and *c* scores from the matching task...137

Table 4.5: Summary of mean % correct responses, *d′* and *c* scores from the matching task by phonological contrast..138

Table 4.6: Summary of mean % correct responses, *d′* and *c* scores from the matching task by OI condition………...138

Table 4.7: Mean response times from the matching task for each L1 group..........................140

Table 4.8: Mean RT (ms) from the matching task for each L1 group by phonological contrast and OI condition (SDs)..140

Table 4.9: GLMM summary for L1 English matching task accuracy.......................................143

Table 4.10: GLMM summary for L1 Arabic matching task accuracy......................................149

Table 5.1: Overview of post-test questionnaire items relevant to learner perceptions of OI influence and importance when learning new words...162

Table 5.2: Random sample of responses to Q6 by both L1 Arabic and English-speakers....164

Table 5.3: Random sample of responses to Q7-8 by both L1 English and Arabic-speakers.165

Table 5.4: Overview of selection criteria, category definitions and level of abstraction for analysing responses to Q6 and Q8...168
Table 5.5: Examples of category agreement between coders for Q6......................................171
Table 5.6: Coding guidelines for Q6 overall evaluation of OI influence during the task.....180
Table 5.7: Coding guidelines for Q6 specific perceptions of OI influence.............................181
Table 5.8: Overview of Q6 QCA results with number of analysis units for each category (% of total units), by L1 group...186
Table 5.9: Coding guidelines for Q8 beliefs about OI importance...201
Table 5.10: Overview of Q8 QCA results with number of analysis units for each category (% of total units), by L1 group...207
Table 6.1: Sample of key studies investigating pronunciation learning strategies...............231
Table 6.2: Questions from the post-test questionnaire related to strategies..........................234
Table 6.3: Random sample of responses to Q13 by both L1 Arabic and English-speakers..235
Table 6.4: Overview of selection criteria, category definitions and level of abstraction for analysing responses to Q13...237
Table 6.5: Examples of category agreement between coders for Q13....................................238
Table 6.6: Coding guidelines for Q13 and LLS used during the task....................................241
Table 6.7: Overview of Q13 QCA results with number of analysis units for each category (% total units), by L1 group...249
Table 6.8: Overview of mean strategy use frequency during the study (*SDs*)......................263
Table 6.9: Mean frequencies of strategy use by L2 English level of L1 Arabic group (*SDs*) 268
Table 6.10: L1 Arabic strategy PCA after oblimin rotation. Factor loadings presented for both pattern and structure matrices...274
Table 6.11: L1 Arabic strategy PCA component correlation matrix.......................................275
Table 6.12: L1 Arabic labelling and interpretation of the variables and components...........275
Table 6.13: L1 English strategy PCA after oblimin rotation. Factor loadings presented for both pattern and structure matrices...280
Table 6.14: L1 English strategy PCA component correlation matrix......................................281
Table 6.15: L1 English labelling and interpretation of the variables and components.........281
Table 6.16: Multiple hierarchical regression predicting lexical encoding accuracy, as measured by participant mean d' scores..287

Table 6.17: Linear regressions of the relationship between L2 English proficiency and strategy usage, based on PCA factor scores....289
Table 7.1: Summary of cluster comparisons by proficiency test score, matching task accuracy and response time (RT)....300
Table 7.2: Summary of cluster comparisons by factor scores of LLS components....300
Table 7.3: Learner profile 1: AR5450539....305
Table 7.4: Learner profile 2: AR5357676....306
Table 7.5: Learner profile 3: AR5247103....308
Table 7.6: Learner profile 4: AR5730442....309
Table 7.7: Learner profile 5: AR5336601....312
Table 7.8: Learner profile 6: AR5336140....314
Table 7.9: Learner profile 7: AR5336722....316
Table 7.10: Learner profile 8: AR5229722....317
Table 8.1 Summary of research questions and key findings....325

List of figures

Figure 1.1: Relationship between core learner variables in the present study........................19

Figure 2.1: A schematic depiction of different psycholinguistic grain sizes (Ziegler & Goswami, 2005, p. 5)..40

Figure 2.2: Dual-route theory of reading aloud (Coltheart, 2006, p. 9)....................................43

Figure 2.3: The major steps involved in learning to read words described within the framework of a multiple-route model of silent reading (Grainger & Ziegler, 2011, p. 8).......44

Figure 2.4: A schematic representation of a perception–production model of L2 segmental acquisition (Colantoni & Steele, 2008, p. 522)...48

Figure 2.5: Map of Arabic dialects from Procházka (2021, p. 222)..71

Figure 2.6: Example of words deriving from the root /k-t-b/, including semantic connection and patterns of vowels for each lexical item...72

Figure 3.1: Overview of the stimuli, including the /auditory form/, corresponding image and accompanying orthographic input (OI)..95

Figure 3.2: Experimental flow of the study session...100

Figure 3.3: Audio check screens for L1 Arabic and English participants...............................101

Figure 3.4: Example screens from the English proficiency test, including the grammar and vocabulary gap fill, listen and answer, and read and answer screens...................................102

Figure 3.5: Example screens from the English version of the background questionnaire....103

Figure 3.6: Word learning task presentation trial sequence..104

Figure 3.7: Word learning task testing trial sequence...105

Figure 3.8: Example of a match and mismatch trial in the audio-visual matching task.......106

Figure 3.9: Distraction check screens for L1 Arabic and English participants......................107

Figure 4.1: Word learning phase presentation trial sequence...128

Figure 4.2: Word learning phase testing trial sequence..128

Figure 4.3: Example match and mismatch trial screens..129

Figure 4.4: Word learning raw accuracy and mean response time data................................131

Figure 4.5: Histogram of response time data for the word learning phase............................133

Figure 4.6: Matching task correct answers for each L1 group by contrast and OI condition139

Figure 4.7: Matching task raw RT by contrast and OI condition for each L1 group.............141

Figure 4.8: L1 English accuracy mixed-model effects, including odds ratios, confidence intervals and level of significance..144
Figure 4.9: Model estimated marginal means for L1 English accuracy, plotted by (a) OI and (b) phonological contrast for both match and mismatch trials...145
Figure 4.10: Model estimated marginal means for L1 English accuracy, plotting the interaction between phonological contrast and OI for match and mismatch trials...............146
Figure 4.11: Spectrograms of the /f-v/ minimal pair items taught with Arabic OI, as an example of partial word-initial devoicing in (b) /vædət/...147
Figure 4.12: L1 Arabic accuracy mixed-model effects, including odds ratios, confidence intervals and level of significance..150
Figure 4.13: Model estimated marginal means for L1 Arabic accuracy, plotted by (a) OI and (b) phonological contrast for both match and mismatch trials...151
Figure 4.14: Model estimated marginal means for L1 Arabic accuracy, plotting the interaction between phonological contrast and OI for both match and mismatch trials......152
Figure 4.15: Model predicted probabilities for L1 Arabic accuracy, plotting trial type by L2 English proficiency..153
Figure 4.16: Model predicted probabilities for L1 Arabic accuracy, plotting the interaction between phonological contrast and OI by L2 English proficiency..153
Figure 5.1: Steps of inductive category development (Mayring, 2014, p. 80).........................167
Figure 5.2: Questionnaire responses for when OI was first noticed during the learning phase of the experiment, by L1 group..174
Figure 5.3: Questionnaire responses for overall spelling preference, by L1 group................175
Figure 5.4: Questionnaire responses for perceived influence of (a) English OI, (b) Arabic OI, and (c) no OI, by L1 group...176
Figure 5.5: Questionnaire responses for overall spelling preference, by proficiency score..177
Figure 5.6: Questionnaire responses for perceived influence (a) English OI, (b) Arabic OI, and (c) no OI, by proficiency score...177
Figure 5.7: Hierarchy diagram showing the relationship between the general types of responses, the main categories, and subcategories for coding responses to Q6.....................179
Figure 5.8: Treemap of L1 Arabic group's perceived influence of OI during the study.......196
Figure 5.9: Treemap of L1 English group's perceived influence of OI during the study......196

Figure 5.10: Questionnaire responses for Q7, by L1 group..199
Figure 5.11: Hierarchy diagram showing the relationship between the general types of responses, the main categories and subcategories for coding responses to Q8.....................200
Figure 5.12: Treemap of L1 Arabic beliefs about why early exposure to OI is important.....221
Figure 5.13: Treemap of L1 English beliefs about why early exposure to OI is important...221
Figure 6.1: Hierarchy diagram showing the relationship between the general types of strategy, the main categories, and subcategories in response to Q13......................................240
Figure 6.2: Notes made by one L1 English participant during the experiment......................248
Figure 6.3: Overview of target words, with accompanying images and written form..........252
Figure 6.4: Treemap of L1 Arabic group's self-reported strategies during the study............260
Figure 6.5: Treemap of L1 English group's self-reported strategies during the study..........260
Figure 6.6: Boxplots for frequency of strategy use (scale 1-5) related to Arabic and English spelling, by L1 group...264
Figure 6.7: Boxplots for frequency of cognitive strategy use (scale 1-5), by L1 group...........266
Figure 6.8: Boxplots for frequency of metacognitive strategy use (scale 1-5), by L1 group..266
Figure 6.9: Boxplots for frequency of strategy use (scale 1-5) related to Arabic and English spelling, by proficiency...267
Figure 6.10: Boxplots for frequency of strategy use (scale 1-5) related to cognitive strategies, by proficiency...269
Figure 6.11: Boxplots for frequency of strategy use (scale 1-5) related to metacognitive strategies, by proficiency..270
Figure 6.12: Inter-item correlations for the L1 Arabic strategy inventory..............................272
Figure 6.13: Scree plot to assess component extraction for the L1 Arabic data.....................272
Figure 6.14: Inter-item correlations for the L1 English strategy inventory............................278
Figure 6.15: Scree plot to assess component extraction for the L1 English data...................278
Figure 6.16: Correlations between learner characteristics, strategies, and accuracy.............285
Figure 6.17: Correlations between factor scores for the LLS components and mean *d′* scores in the matching task...286
Figure 7.1: Indices to guide optimal cluster decision, namely (a) the elbow method, (b) silhouette method, and (c) gap statistic...298
Figure 7.2: Scatterplots for 2 and 3-cluster solutions...298

Figure 7.3: Interquartile range, mean and 95% CIs of proficiency scores, by cluster.............299
Figure 7.4: Bar plots of questionnaire responses related to OI importance and spelling preference, by proficiency cluster...301
Figure 7.5: Bar plots of questionnaire responses related to OI condition difficulty, by proficiency cluster..301
Figure 8.1 Relationship between core learner variables in the present study.........................324
Figure 8.2 Predictions of minimal facilitation of orthography for L2 phonology..................339
Figure 8.3: Predictions of minimal facilitation of cross-scriptal orthography for L2 phonology...341

Acknowledgements

These pages are the result of faith and community. A personal faith that has shifted in shape over my lifetime, but has been formative in my drive to be curious, challenge assumptions and advocate for those whose voices are not often heard. Also, a community that has had consistent faith in the value of my work and my ability to do it well. This has been invaluable in the (many) moments where I lost that perspective for myself. While I am grateful to more than I can name here, those mentioned have been particularly significant in helping me grow seeds of ideas into the chapters that occupy this volume.

I am certain that without the support and guidance of my undergraduate supervisor, Prof. Martha Young-Scholten, I would not have taken the path that I have. Thank you for introducing me to topics that continue to fascinate me and for consistently pushing me forward in my academic journey. Thank you to my PhD supervisors, Prof. Leah Roberts and Prof. Sam Hellmuth, who have helped me grow in so many ways over the course of this project. Thank you for giving me the space to trust my own instincts and say yes to all the additional opportunities that opened up over the years.

This thesis would not have been possible without the support of friends and colleagues, who have traversed their doctoral studies alongside me. Particular thanks to Dr. Yara Abdelaty, Dr. Latifa Alkuwaiz, Dr. Aljawharah Alzamil and Haifa Algothami for the time and expertise they gave to audio recordings, checking translations, and general discussion. Additional thanks to Dr. Mary Barker, Dr. Kate Mooney, Dr. Sophie Moxon, Dr. Sophie Thompson-Lee, Dr. Justyna MacKiewicz, Dr. Carina Mueller, Dr. Şermin Arslan, Dr. Annie Evans, Jessie Shepherd, Junlan Pan and Michelle Hunter, for sometimes talking about research but mostly sharing the ups and downs of life along the way. This also applies to the excellent friendships with housemates in Alma Terrace and Nunmill street.

Last, but by no means least, I am grateful to my family and friends, for their boundless support, throughout my life, but particularly over the last few years. Thank you to my Mum and Dad, and my brothers: David, Sam and Joel, for celebrating every small step forward and helping me laugh (at myself). Thank you to Kristina Lupton and Mirjam Widmann for being so present in my life, despite being separated by both land and sea. Thank you to my partner, Matthew Eberhardt, for more than words could possibly say.

Author's declaration

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for a degree or other qualification at this University or elsewhere. All sources are acknowledged as references. All data collection and analysis were carried out by the author, Louise Shepperd.

Material from this thesis has been published in:
Shepperd, L. (2022). Including underrepresented language learners in SLA research: A case study and considerations for internet-based methods. *Research Methods in Applied Linguistics*, 1(3), 100031.

Chapter 1: Introduction

> It was called the soft g - a 'g' with a little squiggle on top, like this: ğ. Always it had to be preceded by a vowel, and even though it sometimes lengthened the sound of the vowel, it did not have a voice of its own. Every other letter made a distinctive sound, expressed itself loud and clear, except this one. The soft g did not talk. It did not complain or articulate opinions or demand anything. With its puzzling silence and slightly distracted manners it immediately stood out amid the gushing, garrulous letters. It must be a foreigner, I thought. An outsider. An alphabetical outcast. No word in my mother tongue started with it, which I found rather unfair. It was almost as if it was invisible. If you encountered it in the middle of a word, you were supposed to pretend not to have seen it. Just move on and gently skip over it. So the soft g remained mute no matter what the text or context. Yet the more attention I paid to this mystifying letter the more I came to believe that it was trying to tell me something. Perhaps it did speak after all, in its own way, but no one was interested in hearing what it was saying. And somehow my seven-year-old brain associated this unwanted letter with my unwanted left hand. They were both unpopular in the classroom, that's how it felt. Maybe they could connect. (Shafak, 2020, pp. 12–13)

In her book, *How to Stay Sane in an Age of Division*, the novelist Elif Shafak briefly shares her experiences of learning to write in Turkish, as an illustration of navigating alienation and belonging. In the excerpt above, she eloquently explores the relationship between sounds, symbols and herself, with particular reference to the opaque nature of the letter <ğ> (Ünal-Logacev et al., 2019). This anecdote offers insight into the sensemaking of written language patterns, especially when letters do not transparently map to the sounds of a language. Note that it is the written form that is unreliable here, as we are clear about how our first language(s) are spoken and when this is not well-reflected in the written representation. This may seem obvious, but worth remembering when turning to the role of written language in second language learning. Shafak also describes the additional processing required when spelling does not consistently represent sounds, such as the need to suppress attention to the visual form. While this differs from the focus of the present study, in that there are not multiple languages and writing systems to navigate, it echoes thoughts shared by those who participated. Furthermore, this story highlights the ways that people make sense of written and spoken language development through broader experiences of learning. The author, much like the participants in this study, forms idiosyncratic connections between sounds, letters and something that was meaningful to her. While she learned to read early, she struggled to progress with writing, as her left-handedness was not permitted in the classroom. Writing with her left hand became a private pastime, while she practised and gradually improved writing with her "right and respectable hand" (Shafak, 2020, p. 13). She encountered and internalised certain beliefs which then shaped her approach to future learning. By reflecting on these beliefs and playing with inconsistencies, such as the soft-g, she was able to personalise and ultimately progress her learning.

This illustration demonstrates the importance of investigating language acquisition and the influence of literacy from multiple angles. The insight gained from comparative performance alone paints a partial picture of the influences on language processing and development. Without listening to the author's experiences, false assumptions may easily be arrived at about literacy skills and linguistic expression. The present study similarly seeks to explore the topic of orthographic influence on second language phonological acquisition from multiple angles; incorporating participant perspectives with other measures of language learning for a more complete picture of the interwoven factors which pertain to the research topic at hand. This chapter introduces the study with a brief outline of the motivation, research context and structure of the thesis.

1.1 Motivation and contribution

The primacy of spoken input over written input is common across most if not all experiences of first language (L1) acquisition, whereas written input constitutes a large proportion of early second language (L2) input, especially in classroom settings. Thus, exposure to written forms occurs during phonological development in the new language, with arguably far-stretching implications. Phonology refers to the system and categorical organisation of speech sounds in a language, as well as the study of sound patterns across languages. The written representation of these speech sounds in a particular language is what is referred to as orthography. Over the last two decades, there has been an increase in studies examining the influence of orthographic input[1] on L2 phonological development. From this research some argue that visual representations support L2 word recognition, particularly helping to disambiguate confusable sounds that are not found in the known language(s) of the learner (i.e., nonnative phonological contrasts). Alternatively, others present evidence that visual analysis can negatively influence mental representations of L2 phonology, exhibited in non-target perception and production.

Thus far, studies examining the relationship between L1 and L2 orthographic knowledge and phonological development have focused on adult acquisition of nonnative contrasts that are easily assimilated to an existing L1 sound category. Most studies are also conducted with languages where the L1 and L2 share a writing system, predominantly the Roman alphabet. The present study extends this research by teaching words that (1) contain contrasts with different assimilation patterns between L1 and L2, (2) are taught with written input in different scripts, as well as (3) inviting participants to reflect on their language learning perceptions and strategies in relation to written input. The aim is to investigate the influence of orthography on participants' performance learning novel words differing by difficult L2 phonological contrasts, their perceptions about their learning, and the language learning strategies they use to accomplish the task (figure 1.1).

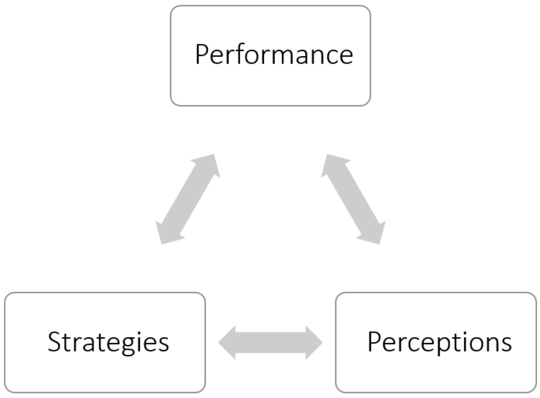


Figure 1.1: Relationship between core learner variables in the present study

Specifically, this study seeks to extend work investigating the orthographic effects of cross-script input, in this case Arabic and Roman scripts. L1 Arabic learners of L2 English, who were literate in both writing systems, with varying L2 English ability, were chosen to

[1] Orthographic input and written input are used interchangeably throughout.

increase understanding of biscriptal literacy, proficiency and other individual differences that may mediate orthographic influence on L2 phonological acquisition. Furthermore, the mixed methods approach to this topic provokes both methodological and theoretical discussion, with directions presented for future research and pedagogical applications. In summary, this work brings together a range of theory and methods to comprehensively investigate the influence of different orthographic forms on phonological acquisition during novel word learning, by adult learners with varying language and literacy experience. This is socially and pedagogically motivated by the impact of migration patterns on L2 classroom diversity, as well as the furthering of L2 acquisition theory and methods in relation to phonology and literacy development.

1.2 Thesis outline

Chapter 2 presents the relevant literature, divided into two main sections: the role of input in second language acquisition (SLA), and the role of literacy and orthography in phonological development. Subsections provide theoretical context and definitions for key concepts applied throughout the thesis. Subsequently, a comparative overview is provided for English and Arabic phonological and orthographic systems, followed by the rationale and research questions for the present study. Chapter 3 outlines the methodology for the study, including stimuli design, procedure and considerations related to internet-based methods. Chapter 4 presents the quantitative analysis and results from the word learning and audio-visual matching tasks. Chapter 5 presents the mixed methods analysis of participant perspectives on the influence of written input during the experiment and on word learning more broadly, emphasising qualitative findings. Chapter 6 presents the mixed methods analysis of self-reported language learning strategies and how they relate to task performance. Chapter 7 then draws together findings on the role of proficiency reported throughout the thesis, through cluster analysis and learner profiles. Chapter 8 discusses the results from chapters 4-7 in relation to the research questions, highlighting the theoretical contributions, pedagogical applications, and methodological considerations, as well as future directions for research. Finally, chapter 9 presents the conclusions that can be drawn from the thesis.

Chapter 2: Literature review

In order to provide context and definitions for key terms, the relevant literature has been divided into two main sections. The first provides an overview of the role of input in second language (L2) acquisition. Seminal theories are detailed, setting the stage for the predominantly cognitive perspectives that are drawn on throughout this thesis. Key concepts are then defined in relation to the characteristics of input and language learners which have been found to play a role in L2 processing and learning. Additionally, context is provided for existing research into learner awareness and strategies in L2 learning. The second section then narrows its focus to orthographic input and L2 phonological development. Attention is paid to literacy experience in different writing systems and how that relates to the processing of spoken and written language. This is followed by the presentation of dominant models accounting for crosslinguistic influences and the anticipated difficulties surrounding the perception and production of sounds that are not found in a learner's first/known language(s) (L1). Empirical studies that closely pertain to the current study are then presented in more detail, providing context for the theoretical and methodological choices that underpin this thesis. Finally, Arabic and English phonological and orthographic systems are detailed, providing background information for the rationale and research questions for the present study.

2.1 The role of input in second language acquisition

One of the central assumptions of the present study is that both exposure to and experience with language influences L2 learning. This is uncontroversial, as exposure to input, understood as incoming linguistic data attended to for meaning, is widely agreed to be essential for language learning (Mitchell et al., 2013). Second language acquisition (SLA) research tends to investigate how input gives rise to learning and the extent to which differences in quantity and quality of input shape language development. These studies often fall within two broad theoretical frameworks. First, nativist or generative approaches assume that humans are born with an innate language faculty, which contains a universal set of linguistic principles and parameters. Input then triggers language-specific settings that constrain acceptable linguistic representations (White, 2014). In contrast, usage-based or emergentist approaches posit that languages are learned in the same way as we learn everything else, and draw on general associative and cognitive processes (N. C. Ellis, 2006; N. C. Ellis & Wulff, 2014). From this perspective, acquisition is driven by regularities and probabilities of occurrence in input, where learners form generalisations involving associations between form and meaning.

While both perspectives agree that input is necessary, the sufficiency of input is a critical point of controversy. Nativists argue that the unconscious knowledge of what is ungrammatical and the ability to produce language that has never been heard is evidence that we know more about our languages than we have been exposed to based on input alone. Additionally, if input was sufficient, why is L2 acquisition seemingly incomplete for many language learners? However, some learners do develop nativelike L2 proficiency, and others do not necessarily pursue nativelikeness as the goal of language acquisition. Usage-based approaches take the position that input is sufficient and emphasise human cognitive capacity to detect distributional information in communicatively-rich social environments, resulting in a system of complex and connected linguistic representations. The reason why ultimate attainment is so variable, if input is assumed to be sufficient, is a continuing point of interest.

A good starting point to contextualise the development of input-related theories is the "Input Hypothesis", presented as part of Krashen's Monitor Theory (Krashen, 1977, 1981, 1985). Krashen distinguished between input and intake (Corder, 1967), noting that we do not acquire all the language we are exposed to and we do not learn all the language we are taught. Note that Krashen also advocates for a distinction between learning and acquisition, which is returned to momentarily. Critically, he argued linguists and language teachers need to consider materials and contexts which do not only provide input but promote intake. The primary requirement for intake was understandable and communicatively meaningful input. This led to the concept of 'comprehensible input', which was ideally a complexity of language just beyond a learners' current level, so as to encourage developmental progress. Another aspect of Krashen's Monitor theory was the 'affective filter hypothesis', which stated that learners differ in their receptiveness to comprehensible input and that emotions play an important role in SLA. For example, affective variables, such as self-esteem and self-confidence, filter input such that less input will reach the "Language

Acquisition Device" (LAD) for those with a high affective filter. Whilst a number of the claims from Krashen's Monitor Theory have been criticised, often based on the difficulty operationalising concepts and testing predictions, ideas underlying this theory continue to ripple out into current SLA theory (Mitchell et al., 2013).

One example is the aforementioned distinction between learning and acquisition, where the former refers to conscious processes and knowledge about language, whereas the later refers to subconscious processes in response to natural language in communicative contexts, which were said to resemble the same processes as used in L1 acquisition (Krashen, 1981). Krashen argued that learned and acquired knowledge could not be eventually integrated. During the same period, Bialystok (1978) also drew a distinction between conscious (explicit) and subconscious (implicit) knowledge in SLA, but claimed that they both interacted with each other. Research into explicit and implicit knowledge has radically expanded since that time and the relevance of these debates to the processing and storage of input in SLA is discussed further in section 2.1.3.1[2].

The 1980s witnessed a torrent of theoretical and empirical research in SLA, including responses to influential work from the 1970s, like that of Krashen. In particular, criticism of Krashen's work arose from cognitive psychology and information processing models of SLA. For example, McLauglin (1987) took ideas from general psychology, regarding memory stores and short-term vs. long-term memory, arguing that learning involved controlled processing which gradually became automatic processing of language. A related concept is that of declarative and procedural knowledge, where declarative knowledge is an object of thought or piece of information that a person is aware of knowing. Meanwhile procedural knowledge relates to the performance of cognitive actions and skills, which can then be automatised. Since the 1980s, understanding around the role of input in SLA has developed greatly. The subsequent sections detail these developments with a focus on the consequences of different factors related to input, the processing and storage of knowledge in the minds of language learners, and distinctions between conscious and subconscious language learning processes.

---

[2] The present study adopts a predominantly cognitive approach, giving greater attention to evidence from this perspective. Additionally, the terms *acquisition* and *learning* are used interchangeably, while *explicit* and *implicit* are used to articulate conscious and subconscious distinctions.

2.1.1 Input-related factors

Particular characteristics of input have been widely researched and used to predict cross-linguistic influence, developmental patterns, ultimate attainment, and learner variability. These characteristics have also been used to compare differences in exposure to input by children and adult language learners. For example, "Children are exposed to vast amounts of auditory native speaker input from birth"(Colantoni et al., 2015, p. 7), or arguably while in utero (May et al., 2011; Vouloumanos & Werker, 2007), whereas both quantity and quality of L2 input is often far more restricted. This may differ in naturalistic settings; however, in the adult language classroom, learners may be limited to a few hours of input per week, with a single speaker to model the language. It could also be argued that the teacher's language is not always input, as language is often modelled for imitation and instruction rather than communication. Instead, learners are likely to listen to and interact with fellow classmates to a large extent. All in all, in both instructed and immersive settings, it is difficult to know how much input L2 learners are exposed to or how much is required for language learning (Piske & Young-Scholten, 2008).

Another critical distinction between child and adult language acquisition is that of modality of input, where visual and auditory senses are considered different perceptual modes. Comprehension and production of written texts do not emerge from ambient exposure and general development, in the way that spoken language does, but rather relies on explicit instruction. This is evidenced by the millions of people around the world who are illiterate, yet are proficient in one or, often, multiple languages (Huettig & Mishra, 2014; Tarone & Bigelow, 2005). Thus, the influence of written input is dependent on literacy. The importance of this point is that at the earliest stages of L1 development the modality of input is predominantly, if not exclusively, spoken language. In contrast, many L2 classrooms assume either (a) pre-existing L1 literacy, and/or (b) the necessity of L2 literacy for language learning. Thus, a substantial proportion of early exposure to language is written, which is often presented simultaneously with spoken input (Young-Scholten, 2002). The early exposure and heavy reliance on written input in adult language learning is of central interest to the present study, and evidence relating to its influence is further detailed in section 2.2.3. Additionally, the ways in which literacy acquisition relates to phonological development is presented in section 2.2.1.

Beyond quantity, quality and modality, there are four further input-related concepts that are essential to consider, namely: frequency, salience, function, and the interactions between each of them. Frequency refers the occurrence of a token or type of item in the input, where high type frequency (e.g., English past tense *-ed*) is more likely lead to the formation of a general category, while high token frequency (e.g., English irregular past *swim/swam*) promotes entrenchment of a particular lexical item. Perceptual salience is the extent to which a linguistic feature is prominent or easy to hear, where low salience cues are more difficult to learn. Then, function relates to the (1) prototypicality of meaning, which depends on the extent to which a particular exemplar corresponds to a defined category, and (2) redundancy, where cues that are unnecessary in order to interpret meaning are not readily learned. The mapping between salience of form and function is also an important factor, highlighted by associative learning theories and referred to as *contingency learning* (N. C. Ellis & Collins, 2009). These concepts are further developed in the next section, which details theoretical accounts of input processing and the learning mechanisms involved.

2.1.2 Processing, comprehending, and learning

Thus far, the aspects relevant to understanding what we mean by input and how it can vary have been discussed. This section returns to the concept of 'intake' and what is actually learned from the input by language learners. A definition of intake has been offered by VanPatten et al. (2020), as "the subset of the input for which a learner can connect form and meaning during real-time comprehension". While there has been consistent interest in understanding the role of input and processing of intake, there remains a lack of consensus about the mechanisms involved in the transition from input to intake. Some scholars have looked to input-related factors, such as frequency of occurrence in the input, or reliability of form-function mappings (N. C. Ellis & Collins, 2009). Just as higher frequency occurrence in the input is related to increased intake, "associative learning theory demonstrates that the more reliable the mapping between a cue and its outcome, the more readily it is learned"(N. C. Ellis & Wulff, 2014, p. 80). These ideas overlap with those set out in the Competition Model (MacWhinney, 2002, 2008), where language learning is driven by the validity, detectability and reliability of cues in the input. Additionally, cue strength is related to competition and expectation, based on the L1. Another related theory is that of Input

Processing (VanPatten, 2014), which asserts that learners need to make correct form-meaning connections as part of comprehension, in order to acquire language. Each of these theories is presented in more detail, to better understand what is involved in the processing of input.

The Competition Model was another theoretical framework to emerge in the 1980s. It combined concepts of cue reliability and availability with cognitive neuroscience and social context to argue for data-driven language learning processes, in opposition to generative positions (MacWhinney, 1987, 2002). Competition is a common construct across psychological theories, exemplified by the well-known Stroop-effect (Stroop, 1935), where a word name competes with the colour in which it is written. When two opposing cues are in competition with each other, MacWhinney posits that the outcome is determined by cue strength. The strength of the cue then depends on cue reliability, availability, cost and detectability. This model was revised to become the Unified Competition Model (UCM) (MacWhinney, 2008, 2012, 2018) of first and second language acquisition, arguing that both L1 and L2 learning were not fundamentally different and social context was central to competition. The UCM included the concept of entrenchment, where neural networks become committed to specific linguistic patterns, especially in early childhood. These existing patterns can then disrupt later L2 development, particularly auditory phonology (Kuhl et al., 2005) and articulatory phonology (Major, 1987).

According to the UCM, entrenchment can be counteracted by resonance, which offers new encoding of L2 patterns through consolidation of memories and reactivation of retrieval pathways[3]. Other pairs of risk and support factors for SLA are outlined in table 2.1. These factors demonstrate the influence of the L1 on L2 learning and processing, where connecting pathways based on the L1 are persistent. For example, whenever a match is perceived between items in the L1 and L2, transfer is expected from the L1. In particular, transfer of auditory and articulatory patterns are noted, and "imperfect transfer" is expected where mappings are similar but not identical, resulting in a foreign accent (further discussion in section 2.2.2). The UCM argues that support factors can help to overcome risk

[3] MacWhinney also states that, "orthography provides a major support for resonance in L2 learning…[However] when the L2 learner is illiterate, or when the L2 orthography is unlike the L1 orthography, this backup orthographic system is not available to support resonance" (MacWhinney, 2012, p. 12)

factors through the gradual establishment of proceduralisation, use of the L2 in inner speech, and active participation in the L2 speech community.

Table 2.1: Unified Competition Model risk and support factors for second language learning (MacWhinney, 2012, p. 34)

Risk factors	**Support factors**
Entrenchment	Resonance
Misconnection	Proceduralisation
Parasitism	Internalisation
Negative Transfer	Positive transfer
Isolation	Participation

As mentioned, several of these constructs are found within another emergentist perspective, namely Associative Learning Theory (N. C. Ellis, 2006, 2008). This framework emphasises frequency-based probabilities in input and introduces concepts of learned attention and blocking. Implicit and explicit learning are also highlighted in this theory, which is true of UCM but to a lesser extent. According to Ellis, learning involves symbolic units, or "constructions", such as lexical items, formulae/chunks and open abstract schemata e.g. [noun stem + plural]. Language representations are then tuned based on frequency, recency and context of specific constructions. Learners also implicitly abstract statistical regularities from groups of constructions, which emerge out of a learner's experience with the language. As Ellis says, "we are conscious of communicating rather than of counting"(N. C. Ellis, 2015, p. 6), but still, we are sensitive to and have knowledge of frequencies, mappings, and transitional dependences.

Frequency in the input is given a central role in language acquisition and associative learning, where processing increases in both speed and accuracy upon more encounters with a cue. However, as mentioned in section 2.1.1, frequency is not the only relevant input-related factor. Salience, reliability, and redundancy of cues also need to be considered when understanding what is learned from the input. For example, if meaning can be deduced without a feature of the input, such as the past tense form of the verb in the sentence "yesterday, I walked home", then the cue is said to be redundant. Furthermore, the content word "yesterday" is more salient than the "-ed", thus overshadowing the past-tense inflection. Ellis argues that overshadowing can then lead to learned attention and attention blocking, where existing associations formed through learned attention to specific cues can then block learning of redundant cues or those with low perceptual salience. In the context of L2 learning, these ideas relate to the influence of associations and learned attention based on L1 experience. Thus, L1 interference can take the form of transferring the content of associations to the L2, but also more broadly biassing attention and processes (N. C. Ellis & Wulff, 2014; Jarvis & Pavlenko, 2008). Ellis (2008, 2015) then argues that learned attention explains where input fails to be converted to intake, which is shaped by L1 entrenchment, and explicit learning is required for L2 learning in order to overcome these implicit biases.

Associative Learning has also been connected to processing theories such as the Efficiency-Driven Processor (O'Grady, 2005, 2008), which details how the capacities and limitations of the human cognitive system shape language acquisition and the drive to avoid high processing costs. Error-driven accounts also propose a learning mechanism driven by prediction and associative learning where "every time an expected outcome is not encountered after a given cue, the strength of its association with that cue diminishes" (Bovolenta & Marsden, 2022, p. 1386; Rescorla & Wagner, 1972). Computational modelling of error-driven language learning in L1 and L2 contexts has been a fruitful area of research in recent years, with much scope for further exploration of learning mechanisms (Bovolenta & Marsden, 2022; Hoppe et al., 2022).

Another processing perspective is offered by VanPatten (1996, 2004, 2014), coming from a generative position. Input Processing Theory is concerned with the early stages of SLA and places central importance on correct form-meaning connections during comprehension. A number of claims are made by the model, such as (1) L2 learners are driven to make lexically-based form-meaning connections, (2) L2 real-time comprehension is demanding in terms of both cognitive processing and working memory, (3) learners have limited capacity for processing and storage of information, and processing is more restricted than for L1 speakers, and (4) L2 Learners draw on universal principles and strategies, as well as transferring L1 processes. In line with these claims, VanPatten argues that forms can only be acquired once processing for meaning is automatic, due to limited attentional capacity. VanPatten also developed an accompanying pedagogical application of this theory, namely Processing Instruction (VanPatten, 2004, 2015). In defining this approach, VanPatten draws attention to the distinction between processing and noticing, where processing instruction is not predicated on awareness in the way that other consciousness raising forms of explicit instruction might be. This further highlights the necessity to understand awareness and conscious attention in language learning, which is the topic of the next section.

To briefly summarise, these three approaches demonstrate that across the generative and emergentist positions, there is agreement that understanding the processing of input is important to unlock what is learned from input and why. These theories differ in their assumptions of internal language learning mechanisms and how they relate to cognitive processes. However, they all acknowledge that language learner processing and memory capacity is limited, so learners necessarily prioritise aspects of the input during processing, which are directed by factors such as cue strength, learned attention, L1 transfer, and a drive to interpret meaning.

2.1.3 The role of consciousness in relation to input

As has been mentioned several times already, the role of consciousness in SLA has been a consistent point of discussion in relation to input processing and representation. The following subsections provide discussion of (1) implicit and explicit distinctions, (2) the nuances surrounding concepts of attention, awareness and noticing, and (3) language learning strategies.

2.1.3.1 Implicit and explicit distinctions

The distinctions between conscious and unconscious knowledge, processes, learning and instruction are regularly referenced in SLA research, where explicit refers to that which is conscious and implicit refers to the un- or subconscious. In an influential paper, Hulstijn (2005) noted that, across the spectrum of generative and emergentist researchers, there is a consensus that L1 acquisition relies predominantly on implicit learning, whereas L2 acquisition depends on both implicit and explicit learning. Understanding the distinction between the two and how they relate to L2 learning is of particular importance for educational contexts, as it is unclear how these constructs can be best exploited in an instructional setting. Hulstijn offers useful definitions for key concepts of implicit and explicit memory, knowledge, learning, instruction, as well as inductive vs. deductive learning, and incidental vs. intentional learning. In relation to the present study, the key concepts are those of implicit vs. explicit *knowledge* and *learning*.

Hulstijn defined explicit and implicit *knowledge* as differing "in the extent to which one has or has not (respectively) an awareness of the regularities underlying the information one has knowledge of, and to what extent one can or cannot (respectively) verbalise these regularities"(2005, p. 130). Here it is possible to see the overlap in the emphasis on verbalisation in definitions of explicit knowledge and declarative knowledge, as mentioned earlier in section 2.1. These types of knowledge are also associated with effortful (explicit) and automatic (implicit) processing. In relation to understanding the role of input in SLA, definitions of implicit and explicit *learning* are particularly pertinent, but also more controversial concepts. Broadly speaking, Hulstijn states:

> Explicit learning is input processing with the conscious intention to find out whether the input information contains regularities and, if so, to work out the concepts and rules with which these regularities can be captured. Implicit learning is input processing without such an intention, taking place unconsciously. (2005, p. 131)

The reasons for controversy and confusion often relate to different understandings about what the object of learning is. For example, Hulstijn focused on regularities of language in relation to explicit learning, whereas Andringa and Rebuschat (2015) have highlighted that both regular and irregular forms can be learned explicitly and implicitly. Andringa and Rebushcat offer clarification by suggesting statistical learning approaches as an example of gradual implicit and automatic learning in response to input exposure. These ideas relate to the accumulation of distributional properties of the input, as mentioned in relation to UCM and associative learning in section 2.1.2. In contrast, explicit learning relates to thinking and talking about a language system, and committing this to memory through practice. It is noteworthy that explicit knowledge is rarely the goal of explicit learning. Instead, instruction promoting explicit learning aims to facilitate L2 input processing and learning of mental representations[4]. For example, this is seen in form-focused instruction (R. Ellis, 2002), processing instruction (VanPatten, 2015), and consciousness raising activities, such as input enhancement (Sharwood Smith, 1993). While it is clear that instruction can affect explicit

[4] Strong emergentist views do not hold that language is symbolically represented in the mind.

knowledge, it is less clear the extent to which it influences implicit knowledge (Robinson et al., 2013).

For clarity on what is widely understood about explicit knowledge, Mitchell, Myles and Marsden summarise the following points of consensus in the field:

> [It] is accessible to conscious awareness, is capable of being put into words, and tends to be used when the participants do not feel under time pressure. It is also thought to be learned faster than implicit knowledge…is learnable at any age, given sufficient cognitive maturity; anxiety reduces the use of explicit knowledge; it is stored as declarative knowledge; it is more inaccurate than implicit knowledge; its quality and use are more prone to individual differences such as working memory capacity; and it is more prone to decay over time than implicit knowledge. (2013, pp. 136–7)

These points add to those already made about the distinction between implicit and explicit knowledge. It also draws attention to the relationship between maturation, memory, and explicit knowledge, connecting to evidence that adults rely more on explicit and declarative learning mechanisms than children (Finn et al., 2016; Zwart et al., 2019). The interface between implicit and explicit knowledge and how that relates to declarative and procedural systems is a point of confusion and controversy (N. C. Ellis, 2015; Suzuki & DeKeyser, 2017). However, broadly speaking, declarative and procedural knowledge pertain to memory systems in the brain, and are well-established concepts in human and animal learning research (Ullman, 2005; Ullman & Lovelett, 2018). Explicit knowledge is understood to only be learned through declarative memory, whereas implicit knowledge can be established through both declarative and procedural systems, but primarily the latter.

It is also relevant that declarative memory likely underpins the encoding, storage, and retrieval of lexical knowledge/the mental lexicon, including phonological and orthographic information. Meanwhile, procedural memory has been linked to sound category learning and the processing of multidimensional perceptual-acoustic cues (Morgan-Short et al., 2014; Quam et al., 2018). An additional useful concept is that of metalinguistic knowledge, also known as metalinguistic awareness, which relates to explicit knowledge and the ability to identify, analyse, manipulate, and reflect on language forms. For example, instinctive abilities to judge whether a sentence is grammatically correct, without the ability to explain why, demonstrate that metalinguistic awareness is not required for implicit knowledge.

2.1.3.2 Awareness, attention, and noticing

Within debates surrounding the role of consciousness and SLA, scholars are interested to know what learners attend to in the input, what level of awareness is involved, and what is the outcome of both attention and awareness. Seminal research on the topic of consciousness and attention was presented by Schmidt (1990), who introduced the Noticing Hypothesis. Building on the wealth of evidence that unconscious processes play an important role in language comprehension and production, Schmidt sought to look more closely at the role of consciousness in language *learning*. Returning to the issue of converting input to intake, Schmidt argued that "intake is the part of the input that the learner notices"(1990, p. 139) and that noticing is required to learn linguistic forms. Additionally, he proposed that paying attention to language forms is broadly facilitative, and may be particularly necessary for the acquisition of redundant features. To clarify these claims, he distinguished three levels of awareness. The first level, perception, was not necessarily conscious, whereas the second level of noticing involved more conscious awareness and related to the subjective experience of stimuli. The third level was said to be understanding, which referred to conscious thinking, problem solving and reflection. He also argued that consciousness does not assume active intent, and intentions may be conscious or unconscious. Consciousness and intention particularly relate to understandings of language learning strategies, which is discussed in section 2.1.3.3.

The construct of noticing has received continued attention in SLA, including Schmidt's modified view that conscious awareness is necessary for the initial encoding of a linguistic feature, or instance of language use, in memory. After which, the strengthening of the representation and organisation within the linguistic system will take place implicitly (Schmidt, 2001). Since then, articulations of the phenomenon have included additional reference to working memory and understanding attention and awareness to be two sides of the same noticing coin (Godfroid et al., 2013). For example, these scholars define noticing as:

> A cognitive process in which the amount of attention paid to a new language element in the input exceeds a critical threshold, which causes the language element to enter

working memory and become the object of further processing (e.g., rehearsal); the traces of this additional processing are stored in long-term memory and, hence, represent intake. (Godfroid et al., 2013, p. 493)

Working memory is understood to be the cognitive mechanisms involved in "the temporary storage, manipulation, and maintenance of task-relevant information during online cognitive operations"(Mitchell et al., 2013, p. 151). It is widely agreed that working memory plays a critical role in comprehension and, as such, is vital for language learning (Baddeley, 2012, 2017). The importance of attentional and memory capacities, and the extent to which this varies between learners, has also been connected to individual differences in L2 learning and outcomes.

While Godfroid et al. (2013) focused on the attentional side of the coin, Andringa and colleagues (Andringa, 2020; Curcic et al., 2019) have sought greater clarity on the subject of awareness. Looking at the emergence of awareness in SLA, Andringa (2020) points to evidence that awareness arises as a product of prior implicit learning and queries whether it is necessary for advancement to higher levels of proficiency. Looking at whether awareness arose spontaneously in uninstructed L2 learning, Andringa found only 33% of participants noticed a target structure in the input, and usage was contingent on awareness. He concluded that "learning a particular structure coincides with the emergence of awareness of that structure"(2020, p. 353) but it remains unclear what triggers awareness, particularly in uninstructed settings.

One of the central issues within the debate surrounding consciousness and language learning relates to how the various constructs can be measured. This is particularly problematic for assessing implicit knowledge, as it is difficult to test what learners do not attend to and are not aware of. In recent years, eye-tracking has proven a useful tool in assessing implicit and automatic processes, offering insight into learning and attention (Andringa, 2020; Godfroid et al., 2013; Roberts & Siyanova-Chanturia, 2013). Additionally, first-person or self-reported data is a common measure of attention and awareness in SLA, supported by the idea that conscious processes and knowledge can be verbalised and acted upon. However, just because something is not mentioned, it does not mean that a learner was unaware or did not attend to it. Equally, depending on the task and time delay, it is easy for learners to forget. As such, these techniques are better placed to assess the relative extent of awareness (Robinson et al., 2013). This developing field will no doubt continue to benefit from methodological and technological advances.

Bringing together the different perspectives outlined above, it is evident that the importance of awareness for language learning is a continuing point of controversy. There is broad agreement that conscious awareness or noticing has a facilitative role in language learning and that both explicit and implicit learning are involved in SLA. What remains unclear is the relationship between input and intake, and the extent to which initial learning depends on awareness. Additionally, questions remain about the influence of existing knowledge on developing awareness and directing attention. Another perspective on awareness and language development is offered by research into language learning strategies, to which the discussion now turns.

2.1.3.3 Language learning strategies

Awareness in language learning does not only refer to metalinguistic awareness of specific rule applications or form-meaning mappings, but also awareness of the learning process itself and conscious strategies applied to support learning. With this in mind, O'Malley, Chamot and colleagues (1987; 1990) presented early research exploring L2 language learning strategies (LLS). They drew on cognitive information processing and skill acquisition theories, the latter of which posits three stages of development: declarative, procedural, and automatic (Anderson, 1982; Anderson et al., 2018). While these developmental stages apply to skill learning broadly, they have been regularly extended to the skills involved in language learning. A key concept, differing from other perspectives on awareness discussed above, is that a learner initially starts from explicit/declarative knowledge *about* a particular skill, which they then act upon to turn the knowledge into proceduralised behaviour of *how to do* something. Once knowledge is proceduralised, practice leads to the automatisation of knowledge, which is generally a lengthier transition than between declarative and procedural knowledge (DeKeyser, 2020). This theory does not specifically comment on the possibility or usefulness of implicit learning, but chooses to focus on the learning of explicitly acquired knowledge. The reason for this focus is related to the application of such theories, because, in language learning, "what matters is fast, accurate, and robust use, the hallmark of automatised procedural knowledge" (DeKeyser, 2020, p. 106). These ideas were applied to LLS by O'Malley and Chamot (1990), who proposed that learners begin with declarative knowledge of a strategy (e.g. 'use context to guess meaning'), which they then put into action and practice until it becomes automatised.

LLS have been a source of controversy over the years, particularly definitions and theoretical underpinnings (Pawlak, 2021). However, there has been a sustained academic and pedagogical interest in the topic, with more recent contributions to the field by Oxford (2016) and Griffiths (2018) seeking to resolve remaining points of contention. For example, Oxford responded to calls from Dörnyei and colleagues (2005; 2003) to abandon the term "strategy" for the all-encompassing idea of "self-regulation", by outlining the interlocking of these concepts in the *Strategic Self-Regulation Model* (Oxford, 2016). Additionally, through a content analysis of existing definitions of LLS, she addressed incoherences in the field and arrived at the following comprehensive definition:

> L2 learning strategies are complex, dynamic thoughts and actions, selected and used by learners with some degree of consciousness in specific contexts in order to regulate multiple aspects of themselves (such as cognitive, emotional, and social) for the purpose of (a) accomplishing language tasks; (b) improving language performance or use; and/or (c) enhancing long-term proficiency. Strategies are mentally guided but may also have physical and therefore observable manifestations. Learners often use strategies flexibly and creatively; combine them in various ways, such as strategy clusters or strategy chains; and orchestrate them to meet learning needs. Strategies are teachable. Learners in their contexts decide which strategies to use. Appropriateness of strategies depends on multiple personal and contextual factors. (Oxford, 2016, p. 48)

This definition draws attention to the consensus that LLS are active, conscious, mental processes, which are goal-oriented and situated in the context of the individual, the task and the learning environment. Oxford (2016) specifies that strategies are employed with the purpose of self-regulation, including cognitive, emotional, and social aspects of learning. In contrast, Griffiths (2018) offers a more condensed definition, intentionally omitting mention of strategies being conscious or related to regulation.

> Language learning strategies are actions (the learner has to DO something); chosen by learners (as distinct from being imposed by someone else, e.g. the teacher); for the

purpose of (they are goal-oriented); learning language (as distinct from e.g. communicating). (Griffiths, 2018, p. 19)

The reason for refining the definition in this way is to capture the gradual automatising of strategies with experience, echoing ideas from skill acquisition theory. Additionally, while it is accepted that regulation is a pivotal aspect of strategy use, Griffiths argues that some strategies, such as cognitive strategies, are not regulatory and regulation is not the primary goal of LLS. Therefore, as language learning is the primary goal of strategy use, regulation should not feature within the core definition. Overall, these accounts both point to greater harmony than disagreement within the field, as well as productive routes forward.

Within these broad definitions, LLS are often classified into metacognitive, cognitive and socio-affective strategies (O'Malley & Chamot, 1990). Strategies in the metacognitive domain include processes such as monitoring, evaluating, and planning of learning, while cognitive strategies refer to processes including analysing, comparing and reasoning. Social and affective strategies facilitate communication within sociocultural contexts and modulate emotional responses to learning, respectively. Classification systems and tools have evolved over time, such as Oxford's (2016) outlining of self-regulation strategies for cognitive, motivational, affective and social domains of L2 learning, as well as overarching metastrategies that reach across multiple domains. Alternatively, Griffiths (2018) focuses on how strategy usage relates to attainment, by categorising both cognitive and metacognitive strategies into *core*, *base* and *plus* strategies. This distinction reflects evidence that advanced students not only use more LLS more frequently, but typically favour the *plus* group of strategies. Meanwhile, *base* strategies are more typical of lower-level students and *core* strategies are found equally among all students. The relationship to attainment is important, as strategies can change with time and training, which underlines the pedagogical relevance of this line of research. This is connected to aims for the present study, where it is of interest to discover whether specific strategies, in relation to orthographic input and phonological acquisition, support learning and could be recommended for instructional settings.

2.1.4 Summary of the role of input in SLA

This section has provided the theoretical context and key terms relevant to the present study in relation to how input is understood in SLA research, with particular focus on cognitive perspectives on input, processing and consciousness. There is broad agreement that language learning is dependent on exposure to input, and that quantity and quality of input differ substantially when comparing L1 and L2 acquisition. Additionally, the role of consciousness in relation to L1 knowledge and learning is said to differ from L2 learning, in that L1 language learning is predominantly implicit, while L2 learning relies on both explicit and implicit learning. While many questions remain around the role of consciousness in language learning, there is a consensus that conscious awareness has a facilitative effect on SLA. This is important to the discussion of input, as it relates to the processing, storage and learning based on input, and how this can be effectively adapted in instructional settings.

There remains a lack of clarity around what is involved in the conversion of input to intake, and how the aspects of the input we attend to go on to shape our learning trajectories. As humans have limited cognitive capacity, in terms of both processing and storage of information, it is logical that certain aspects of the input are prioritised during language processing and representation. Relevant factors to understand this prioritisation relate to both the input itself and the characteristics of the learner. Section 2.1.1 highlighted the importance of cue frequency, salience, and reliability in the input, while section 2.1.2 drew attention to internal mechanisms within language learners which direct processing and attention, based on L1 experience, cue strength, and the drive to connect forms to meaning. Section 2.1.3 went on to detail the role of consciousness in language learning, pointing out continuing uncertainty around the importance of awareness and its place in developmental orders. Turning to the next section of this review, these general concepts related to input and SLA are developed with a specific focus on written input and phonological acquisition.

2.2 The role of literacy and orthography in phonological development

Another central assumption of this thesis is that literacy experience and exposure to orthography in the input influence L2 phonolexical development. This view is more controversial than the broader role of input outlined in the previous section. However, the research presented here demonstrates the connection between literacy and phonology, and reasons why the influence of orthographic input on the lexical encoding of L2 phonology deserves greater consideration. The lack of theory, despite increasing evidence, is also highlighted. This route of enquiry is motivated by the high variability in adult SLA ultimate attainment, where the mastering of foreign sounds is notoriously difficult, in comparison to language(s) learned in childhood. Here, the mastery of L2 sound systems is not related to the goal of nativelikeness, but rather to the importance of L2 phonological acquisition for comprehension and comprehensibility.

Processing and learning based on auditory input can be considered the first port of call for language acquisition (Piske & Young-Scholten, 2008). This is partly due to the fact that, for hearing individuals, our first contact with language is auditory, as well as the early attunement of language specific perception (Werker & Tees, 1984). Related observations were made in the previous section, where 2.1.1 highlighted the differences between the quantity and modality of input typical of L1 and L2 acquisition. Section 2.1.2 went on to note the consequences of entrenchment for L2 phonology, connecting adult L2 learners' extensive knowledge of their L1 to cross-linguistic influence throughout L2 development. Part of that pre-existing L1 knowledge and experience relates to literacy for many adult learners[5]. Literacy is a complex construct to define, and is often explored from two key perspectives. The first focuses on the cognitive consequences of literacy and the processes involved in encoding and decoding text (Goody, 1987; Olson & Torrance, 1991). The second emphasises that literacy encompasses multiple socially-situated practices (Gee, 1991; Gee, 2001; Street, 1984, 1994). In line with Tarone and Bigelow (2005), the present study sees these as complementary perspectives that explore related but different aspects of literacy. As the present study is predominantly concerned with the impact of literacy and written input on mental representation and cognitive processing, references to literacy in this thesis align predominantly with the first perspective.

Literacy is both dependent on and developing of phonological awareness, impacting L1 and L2 perception (Elbro & Pallesen, 2002; Goswami, 2002; Horlyck et al., 2012; Koda, 2007). Motivated by the point made in 2.1.1 it is of interest that, in contrast to L1 acquisition, exposure to L2 written language often does not follow on sequentially from spoken language, but is often presented simultaneously. In fact, literacy skills can be drawn upon to learn new language from written forms before hearing the auditory form. Broadly speaking, written language tends to make up a large proportion of total input in adult L2 instructional settings, including the earliest stages of development. Thus, it is likely that orthographic

[5] Most participants in L2 phonological acquisition research are highly literate (Colantoni et al., 2015)

cues contribute heavily to the development of L2 lexical representations, with far-reaching effects (Bassetti et al., 2015). In order to explore these effects, section 2.2.1 outlines the impact of literacy development on phonological awareness and defines key concepts relevant to understanding literacy in different writing systems. Then section 2.2.2 presents dominant models of L2 speech learning and perception. This provides theoretical context for discussion of L2 phonological development, as well as evidence of the lack of formal attention devoted to orthographic input, to date. Following on from this, section 2.2.3 provides an overview of empirical work focussing on the influence of orthography on L2 phonological acquisition, with particular attention given to (1) lexically encoding nonnative contrasts during word learning, and (2) learning across writing systems.

2.2.1 L1 phonological awareness and literacy development

In chapter 1, phonology was defined as the system or categorical organisation of speech sounds within a particular language, while orthography refers to the written representation of those sounds. Orthography does not fully represent phonology and also represents more than phonology, reflecting linguistic, historical and cultural considerations (Venezky, 2005). As written language is a symbolic representation of the sounds of a language, literacy requires an awareness of the ways in which words can be broken down into smaller phonological units and an understanding of how print encodes the relevant units of sound. This ability to recognise important elements in spoken language and their relations to a writing system when reading is why researchers connect reading to metalinguistic awareness, and phonological awareness in particular (Goswami & Bryant, 2016; Koda, 2007; L. Verhoeven & Perfetti, 2017). Phonological awareness "comprises the ability to recognize, identify, or manipulate any phonological unit within a word, be it phoneme, rime, or syllable" (Ziegler & Goswami, 2005, p. 4) and is a strong predictor of reading acquisition across languages, alongside oral language proficiency (Nation & Snowling, 2004).

The emergence of phonological awareness is reportedly sequential, moving from awareness of large phonological units towards a sensitivity to small phonological units (Anthony & Francis, 2005), following the order presented in figure 2.1. Drawing together evidence from crosslinguistic studies, Ziegler and Goswami (2005) demonstrated that children from all language backgrounds develop structured phonological systems with knowledge of words, syllables, onsets and rimes (large grain sizes). However, phoneme awareness (small grain size) emerges as a result of learning to read. The relationship between phonemic sensitivity and reading instruction, rather than age or developmental stage, is supported by evidence that illiterate adults perform poorly on phoneme awareness tasks, compared to syllable and rime detection tasks, and this improves with literacy instruction (Morais et al., 1979, 1986). Further, phoneme awareness is a better predictor of alphabetic literacy development than other aspects of phonological awareness, such as onset-rime skills (Hulme, 2002). Alphabetic literacy is the key point here, as it appears to be the mapping of individual phonemes to graphemes which predicts the emergence of phonemic awareness, rather than reading acquisition in general.

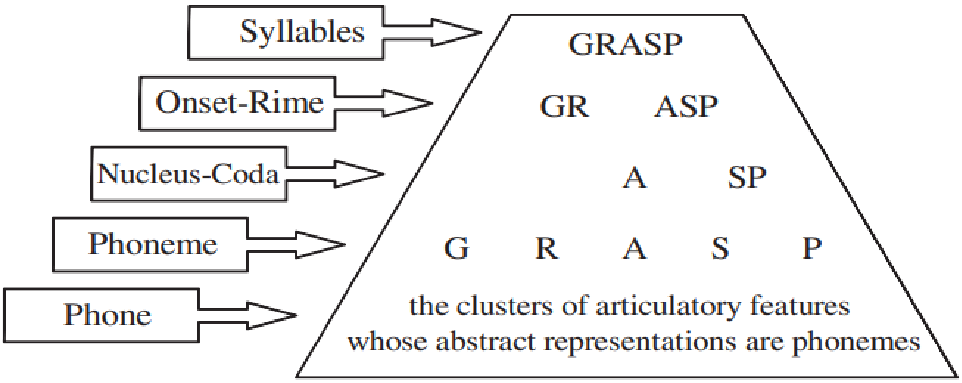


Figure 2.1: A schematic depiction of different psycholinguistic grain sizes (Ziegler & Goswami, 2005, p. 5)

All writing systems encode language in some way, whether at the level of morphemes (logographic), syllables (syllabic), phonemes (alphabetic) or by mixing levels (Ziegler & Goswami, 2005)[6]. Alphabetic languages, such as those that use the Roman, Greek or Cyrillic alphabet, represent individual consonants and vowels with graphemes and vary

[6] Orthographic scripts are not naturally occurring phenomena, but rather relatively recent cultural inventions (Huettig & Mishra, 2014). Written languages reflect historical events, trading, colonial powers, technological advances and linguistic evolutions, amongst other things. However, some argue the relationship between sound and symbol is often not entirely accidental. For example, a logographic system is an efficient visual representation of Chinese morphemes, which are predominantly monosyllabic with a high level of homophony. Equally, the regular moraic structure of polysyllabic words common in Japanese, and the limited number of syllables, lends itself well to a syllabary system. In contrast, Indo-European languages reportedly exhibit less homophony and have more complex syllable structures, making an alphabet more efficient (Katz & Frost, 1992). This, however, is a controversial stance.

by level of consistency. The consistency between phoneme-grapheme correspondences (GPCs) in alphabetic orthographies is described by the Orthographic Depth Hypothesis (Katz & Frost, 1992). An orthography with one-to-one GPCs is considered shallow, while one-to-many GPCs would indicate a deep orthography. For example, English would be an example of a deep orthography (<thought> - /θɔːt/, <though> - /ðəʊ/[7]), whereas Spanish has a shallow orthography ('to think' <pensar> - /pensaɾ/). At this point, it is useful to clarify that *writing system* refers to the broad level of mapping between sounds and symbols in a language, as outlined above, while *orthography* refers to the mapping between symbols and the sounds of a specific language. *Orthographic script* then refers to the symbol system used to represent a language. To illustrate this, English and Spanish are both represented with alphabetic writing systems, and share the Roman alphabetic script. However, the orthographies of the languages differ in their GPCs. Not only do they vary in consistency of representation, but a symbol shared by both orthographies can represent a different sound in each language (e.g., <v> corresponds to /v/ in English but /b/ in Spanish).

Examples of other writing systems include logographic scripts, such as Chinese characters, and syllabaries, like Japanese hiragana and katakana. Additionally, Abjads predominantly represent consonants, and vowels only in some cases, such as those found in Arabic, Hebrew and Persian languages. Then, abugidas or alphasyllabaries, which are common in South and Southeast Asian languages, represent a consonant and vowel together with a single written unit. The Korean hangul script is debatably alphabetic and syllabic, in the sense that individual phonemes are consistently represented with individual graphemes, but these letters are then spatially displayed in units corresponding to syllables. Thus, writing systems not only vary in the level of phonology that they represent, but also mix levels and vary in the consistency of that representation (Verhoeven & Perfetti, 2017). The critical point being that these varying grainsizes and consistency of written representations influence the rate and processes of developing literacy (Goswami & Bryant, 2016; Seymour et al., 2003; Ziegler & Goswami, 2005).

In order to explore and explain the mappings between the symbol and sound systems of a language and reading development, Ziegler and Goswami (2005) presented the Psycholinguistic Grainsize Theory. According to the authors, reading is grounded in

[7] < > denotes orthographic representation and / / denotes phonological representation

phonological processes and "because languages vary in phonological structure and in the consistency with which that phonology is represented in the orthography, there will be developmental differences in the grain size of lexical representations and reading strategies across orthographies"(Ziegler & Goswami, 2005, p. 18). They reported evidence that reading acquisition is faster in shallow orthographies, compared to deep orthographies. For example, a study conducted across 14 European countries by Seymour, Aro and Erskine (2003) showed that children during the first year of reading instruction in a consistent orthography performed close to ceiling with both real and nonwords. Accuracy then reduced in line with the inconsistency of the language, with notably poor performance for those developing literacy in English. With these differences in mind, Ziegler and Goswami argued that the reliance on grapheme-phoneme recoding strategies in consistent orthographies accounts for faster rates and higher accuracy in reading development. In contrast, learners of more inconsistent orthographies cannot rely as much on smaller grain sizes, and therefore develop a variety of strategies that include larger grain sizes, such as rime patterns and whole word recognition.

Additional discussion of reading across different writing systems can be found in Verhoeven and Perfetti (2017). However, Share (2008, 2021) has highlighted the Anglocentric focus of reading acquisition research, where little attention is paid to non-alphabetic languages, and the inconsistencies of sound-spelling mappings in English have received disproportionate levels of attention. The limited consideration of non-alphabetic orthographies and focus on consistency is certainly apparent in both the orthographic depth hypothesis and psycholinguistic grainsize theory. More recent research by Vaid et al. (2022) has demonstrated the need for more research into biscriptal bilingualism, to develop theories in relation to literacy and language learning. For example, they presented evidence that conceptualisations of speech sounds differ in each language of bilingual biscriptals, where a phoneme deletion task revealed Hindi-English bilinguals identify phonemes as the first sound in English words but a syllable for Hindi words. This raises important questions which are beyond the scope of the present study, but underscores the need for research with more diverse participants, languages and writing systems.

The orthographic depth hypothesis (Katz & Frost, 1992) and psycholinguistic grainsize theory (Ziegler & Goswami, 2005) both align with ideas from the dual-route model of word recognition and reading aloud (Coltheart, 2006; Coltheart et al., 2001) (figure 2.2). According to this theory, information about words is stored in the orthographic, phonological, and semantic lexicons. These lexicons each represent and store knowledge about the spellings/visual form, pronunciation, and meanings of words, which then offer different routes for reading. The two routes are the lexical (direct) and non-lexical (indirect) routes, where a familiar word can be recognised and read aloud based on the representation of the word in the mental lexicon, including phonological information. In contrast, an unfamiliar word that has no corresponding representations in the mental lexicon will need to be processed via the non-lexical route (i.e., GPC rules). This choice of route does not only relate to unfamiliar words, but also consistency of GPCs. For example, consistently spelled words can be read via either route, whereas inconsistent spellings can only be correctly read through the lexical route (Coltheart, 2006). While this theory is useful for understanding the mental representation of different types of information and the processes involved in skilled reading, there is limited extension to the development of reading acquisition and the specific constraints that shape orthographic representation.

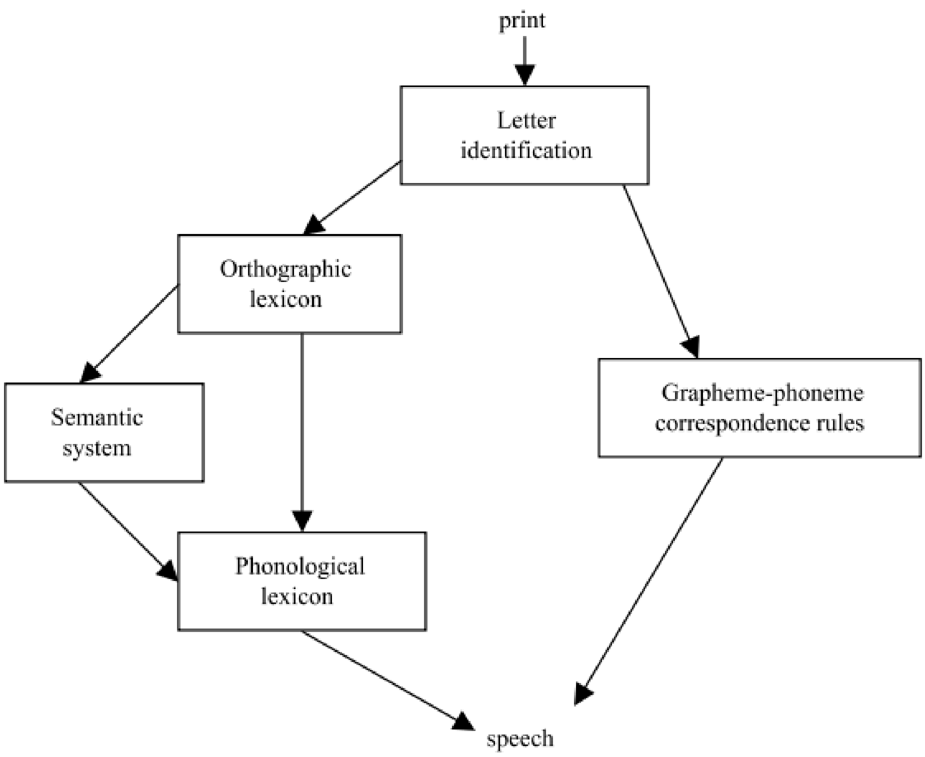


Figure 2.2: Dual-route theory of reading aloud (Coltheart, 2006, p. 9)

In response to these shortcomings, Grainger and Ziegler (2011) proposed a dual-route approach to orthographic processing which integrates the architecture of the bi-modal interactive-activation model (BIAM) of visual word recognition for silent reading (Diependaele et al., 2010; Grainger & Ferrand, 1994; Grainger & Ziegler, 2008). Figure 2.3 outlines the development from a dual-route approach to a multiple-route model for reading acquisition. According to this model, beginner readers (1) start mapping individual letters to individual sounds (phonological recoding). This then develops to (2) parallel independent letter processing and (3) sublexical representations. Within the sublexical representations, (3a) large/coarse-grained representations provide rapid connections to semantics, while (3b) small/fine-grained representations modify processes involving GPC rules and the development of morpho-orthographic representations.

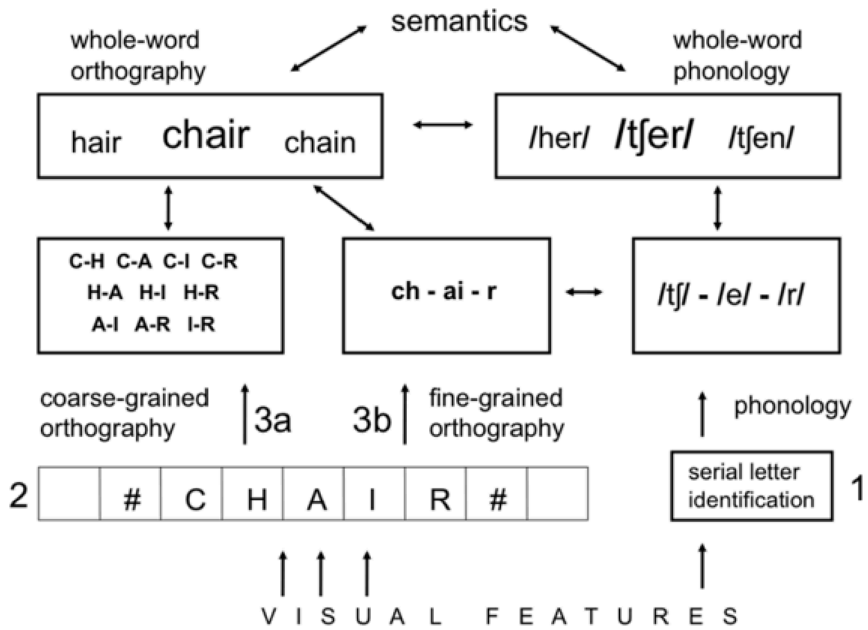


Figure 2.3: The major steps involved in learning to read words described within the framework of a multiple-route model of silent reading (Grainger & Ziegler, 2011, p. 8)

The reason for highlighting these models of skilled and beginner reading is because "once established in one language, reading skills transfer across languages, and are incorporated in learning to read in another language"(Koda, 2013, p. 3). The transferring of reading skills between languages is related to proficiency in each language and the automaticity of reading strategies, where L2 reading can be slowed by the transfer of inappropriate or inefficient reading strategies (Bialystok, 2001). While the present study is not explicitly concerned with L2 reading, the processes involved in L2 spoken and written word recognition are highly relevant. Therefore, it is important to recall ideas of crosslinguistic transfer and the automatic activation of mappings based on rehearsed L1 skills and L2 input, outlined in the first half of this literature review. Specifically, L2 reading skills "evolve through crosslinguistic interaction between transferred L1 skills and emerging L2 orthographic knowledge"(Koda, 2013, p. 3), which is then influenced by L1 and L2 orthographic knowledge and the similarity between the two orthographic systems.

Another influential theoretical framework related to skilled reading is the Lexical Quality Hypothesis (LQH) (Perfetti, 2017; Perfetti & Hart, 2002). According to this hypothesis, each word is a unique combination of three key constituents: (1) orthographic form, (2) phonological form, and (3) meaning. The quality of knowledge of each constituent and the words they form varies between individuals and words. Skilled reading is, then:

> …the synchronous activation of these constituents such that orthographic form, phonological form, and meaning features are perceived not as pieces, but as a single whole. Lexical quality determines the extent to which word reading is one of seamless activation for individual words for a given reader (Perfetti, 2017, p. 52)

Consequently, the quality, stability and accessibility of lexical representations underpin the development of automatised word reading, which then leads to improved processing capacity and fluid comprehension. When comparing high and low skilled readers, the authors suggest that a key difference is the precision of the orthographic form, where less stable knowledge results in longer lasting confusion between competing homophones e.g. rain vs. reign. Other examples are given for encoding underspecified meaning and phonological forms, leading to greater interference from words that appear to have higher similarity due to the lack of precision. In other words, precise knowledge of the constituent parts that make up lexical representations results in improved inhibition of word competitors. The importance of lexical quality for both early L1 and L2 reading development has been demonstrated by Verhoeven and colleagues (2019).

The tight relationship between phonology and orthography is not only evident in visual word recognition, but has far-reaching effects on spoken language processing, as well as other areas of cognition (Huettig et al., 2018; Huettig & Mishra, 2014). For instance, spoken word recognition had been found to be slower for words with inconsistent spelling (Ziegler et al., 2004; Ziegler & Ferrand, 1998), as well as evidence of online activation of orthography during spoken word recognition (Perre & Ziegler, 2008). Additionally, automatic activation of orthography has been found during speech perception, where phonological priming was modulated by orthographic overlap during spoken word processing (Chéreau et al., 2007). In a study yielding similar results, Taft and colleagues concluded that "in literate adults, orthography is important in speech recognition just as phonology is important in reading" (2008, p. 366). There is substantial evidence that literacy fundamentally changes the way that language is mentally represented and how it is processed, irrespective of the modality that language is presented in. This is further underscored by neurological evidence that literacy profoundly alters the brain (Dehaene et al., 2015; Skeide et al., 2017).

In summary, metalinguistic awareness, and phonological awareness in particular, play a central role in literacy development. Returning to the questions raised in section 2.1.3.2, related to the importance of awareness for learning, it appears phonological awareness is a requirement for reading acquisition. However, it is important to state that the relationship between orthographic and phonological knowledge is reciprocal. Phonological awareness predicts literacy development, which in turn results in increased sensitivity to units of phonology, especially phonemes in the case of alphabetic orthographies. The interwoven relationship between phonology and orthography is further exhibited in the automatic activation of orthographic knowledge during spoken language processing. Orthographies all represent spoken language but vary in terms of grainsize and consistency. These differences then influence reading strategies across languages and can be transferred when learning a new language. Several theoretical models explain the routes by which orthographic and phonological input is processed during visual word recognition; the primary claim of which is that there are two main routes, lexical and non-lexical. These two routes underline the importance of the vocabulary size for the former and the consistency of GPCs for the latter. The quality of lexical knowledge is also a relevant consideration.

2.2.2 Models of L2 phonological learning

In order to understand the influence of literacy experience and orthographic input on L2 phonology, it is necessary to first outline the dominant theories that account for the factors and mechanisms involved in L2 spoken language acquisition. These models, in their various ways, are concerned with the challenges that L2 learners face with developing target-like perception and production (Colantoni et al., 2015). As mentioned, foreign accentedness and perceptual limitations are phenomena typically connected to additional languages learned outside of childhood. While it is often not a priority to entirely lose an accent, phonological development must progress with sufficient precision and robustness so that spoken language processing, perception and production result in comprehension. Well-known difficulties include the "rocket" and "locket" distinction for Japanese speakers (Aoyama et al., 2004) or "bet" and "bat" for Dutch speakers (Broersma, 2005). The dominant models of speech learning tend to focus on either perception or production, with varying interest in the different stages of language development. The three models presented below are the Perceptual Assimilation Model (PAM) (Best, 1995; Best & Tyler, 2007; Tyler, 2019), Speech Learning Model (SLM) (Flege, 1995; Flege & Bohn, 2021) and Second Language Perception Model (L2LP) (Escudero, 2005; Escudero & Boersma, 2004; van Leussen & Escudero, 2015). Each model has been revised over the years, and the following discussion focuses on the most recent versions of these models, namely PAM-L2 (Best & Tyler, 2007), SLM-r (Flege & Bohn, 2021) and revised L2LP (van Leussen & Escudero, 2015).

Before discussing the contrasts and commonalities between these accounts, the model of segmental acquisition from Colantoni and Steele (2008) offers a more holistic overview of principles found across these theories, and is relevant to the segmental focus of the present study (figure 2.4). Segmental acquisition refers to the phonological level of individual consonants and vowels, as opposed to prosodic phenomena, such as stress and intonation. Core concepts are illustrated in figure 2.4, including the starting point of input, where segments are compared cross-linguistically based on L1 and L2 phonological knowledge, and then classified in relation to L1 categories. The system of classification in this model incorporates ideas proposed in SLM, which is presented in further detail shortly. This classification system is the deciding factor as to whether a new category needs to be established or whether the segment is perceived to correspond to an existing category in the

L1 phonological inventory. The term 'interlanguage' is then used in the model, indicating that knowledge of the developing L2 falls between that of the L1 and L2, based on crosslinguistic influence (Selinker, 1972). The developing L2 representations feed into articulatory planning, where both L1 articulatory patterns and universal constraints of airstream and articulators shape the spoken 'output'[8].

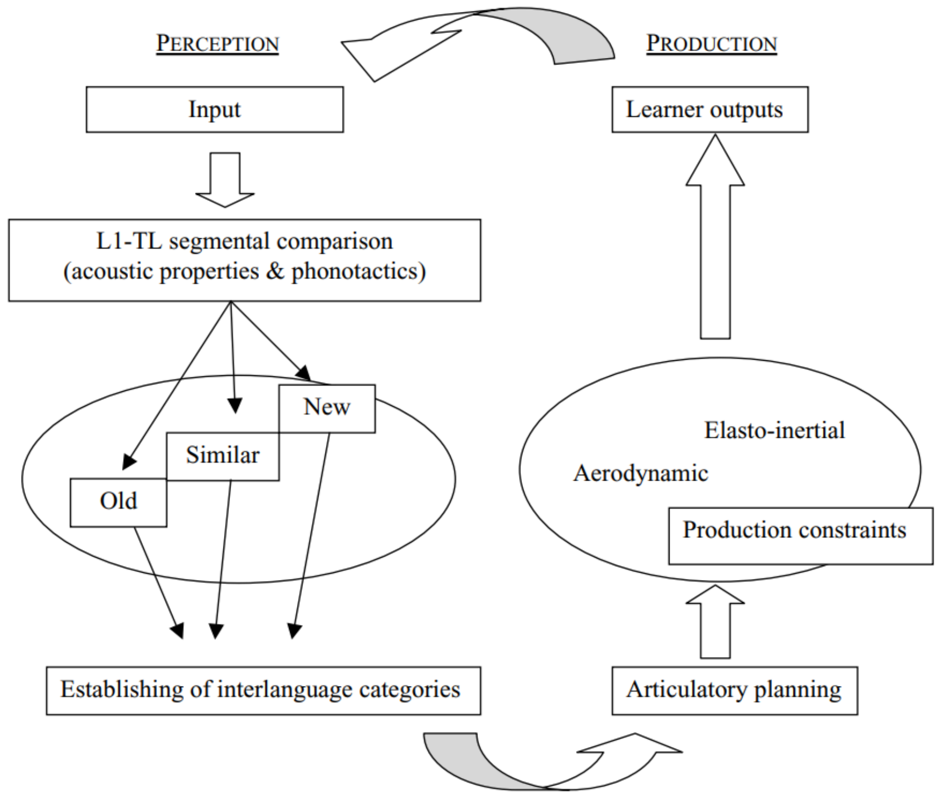


Figure 2.4: A schematic representation of a perception–production model of L2 segmental acquisition (Colantoni & Steele, 2008, p. 522)

With this overview in mind, attention turns to the specifics of the three theoretical models: PAM, SLM and L2LP. These are by no means the only accounts of L2 perception and production. However, they have been chosen based on their substantial and sustained influence on the field, as well as their relevance to the current study. Each of these models emphasises a perceptual basis for L2 speech development and makes predictions about the difficulty of categorising L2 phonological segments based on cross-linguistic influence from L1 phonological categories. However, they then differ in their theoretical assumptions,

[8] Research associated with the Interaction Hypothesis (Long, 1996) and the Output Hypothesis (Swain, 2005; Swain & Lapkin, 1995) argues that output plays an important role in language learning, alongside input (S. M. Gass & Mackey, 2006).

research aims and predictions. The Perceptual Assimilation Model (PAM) (Best, 1995; Best & Tyler, 2007) focuses on naïve listeners and their perception of nonnative speech, taking a direct realist perspective. Accordingly, the authors argue articulatory gestures are directly perceived from the speech signal and then higher order articulatory invariants are detected as part of perceptual learning, in line with an articulatory phonology view (Hall, 2017). They propose that the difficulty of perceiving L2 categories is based on assimilation to L1 categories. Their research was extended from naïve listeners to L2 learners in PAM-L2 (Best & Tyler, 2007), where the phonological level is argued to be more crucial for L2 learners. This is due to their developing interlanguage and active pursuit of meaningful distinctions, in comparison to naïve listeners. The authors additionally highlight differences in attentional focus based on learner proficiency, context and goals.

The principal predictions of PAM relate to different assimilation patterns and associated difficulty, depending on the relationship between L1 and target phonological contrasts. The easiest contrast to acquire would be a two-category assimilation (TC), where each sound within a phonological contrast assimilates clearly to two sounds within an existing contrast in the L1. For example, considering the languages relevant to the present study, /m-n/ is an established phonological contrast in both English and Arabic, thus Arabic learners of English are likely to perceptually discriminate minimally contrasting items with ease (e.g. 'mine' and 'nine'). In contrast, single-category assimilation (SC) is predicted to cause the most problems for perception, which occurs when two L2 categories are perceived to be equivalent to the same L1 category. Returning to the example of Arabic learners of English, /f-v/ is a well-established contrast in English but not Arabic, as the voiced /v/ labiodental fricative is not typically found in the Arabic phonological inventory (see section 2.3.1). Therefore, if /f/ and /v/ are perceived to be equally good/poor examples of /f/ (SC), then the learner will find them difficult to discriminate and minimal pairs differing by this contrast are likely to be perceived as homophones (e.g. 'fine' and 'vine').

Alternatively, a category of goodness assimilation (CG) is possible, which is similar to SC but one category is perceived to be more deviant than the other. In this case, if /f/ is perceived to be a good exemplar and /v/ a poor exemplar, then learners should be able to discriminate these L2 phones well and eventually form a new category for the deviant L2 phone. However, the authors do not state the details of the developmental progression for the formation of this new category. Three additional assimilation patterns are detailed in the model, namely uncategorised-uncategorised, categorised-uncategorised, and non-assimilable contrasts. As these patterns are less pertinent to the present study, further details are not provided here.

As mentioned, PAM-L2 typically forms perception-based predictions for the earliest points in language exposure. In contrast, the Speech Learning Model (SLM) (Flege, 1995) has focused more on L2 production, and the developmental trajectory of language learning across the lifespan, particularly ultimate attainment. Similarly to PAM, SLM proposes a cross-linguistic equivalence classification, where L1 and L2 categories exist in a shared phonological space and L2 speech errors relate to perceptual bias. However, SLM focuses on individual sounds rather than contrasts, and their representation in long-term memory. Additionally, this equivalence classification is based on statistical distributional properties of L2 input. Overall, SLM predicts that, as phonetic dissimilarity increases between L1 and L2 sounds, the easier it will be to perceive a cross-linguistic distinction and form a new target-like category.

Thus, SLM would likely frame /f/ as an *old* or existing sound, therefore requiring no additional learning to perceive and produce. Meanwhile /v/ would be judged as phonetically *similar* to /f/ and a L1-L2 composite phonetic category would develop, rather than the required new category. However, the likelihood of improved perceptual distinction and the formation of a new category is predicted by increased L2 experience. This is claimed to be more difficult for learners than learning a *new* sound, which is perceptually dissimilar from the closest L1 category. SLM-r (Flege & Bohn, 2021), the recently revised version, went on to emphasise age effects and the claim that L2 learners make use of the same mechanisms and processes as L1 children when learning L2 speech. The authors link the differences in outcome to (1) automatic substitution of L2 sound for L1 sounds initially, (2) interference from pre-existing L1 phonetic categories, which can hamper the formation of new L2 categories, and (3) differences in input that form the basis of L2 sounds compared to L1 sounds. Each of these links ties in with broader theories of language learning presented in the first half of this review.

The final model to mention is the Second Language Perception model (L2LP) (van Leussen & Escudero, 2015) which focuses on phonological contrasts in relation to lexical recognition. According to the L2LP, "learners will initially perceive L2 sounds in a manner resembling the production of these same sounds in their L1 environment", arguing that the initial state of L2 learning is, in fact, the outcome of L1 learning. Detected acoustic differences and similarities between the phonemes of the two languages will then shape development. Adding some confusion, L2LP labels the scenario described as SC assimilation by PAM as a *new* scenario. This would apply to /f-v/, as learners must create a new category or split an existing category. Meanwhile, the TC assimilation would be described as a *similar* scenario, where L1 categories are replicated and gradually adjusted to the boundaries of the L2 contrast, which would apply to the /m-n/ contrast here. Therefore, in contrast to SLM, *new* sounds pose more difficulty than *similar* sounds. However, this reflects a difference in terminology rather than direction of predictions, as all three models centre around category assimilation, adjustment and formation.

A critical addition to the L2LP model draws on computational modelling, inspired by connectionist learning frameworks (Gaskell & Marslen-Wilson, 1997; McClelland & Elman, 1986), which overlap with interactive-activation models mentioned in relation to reading (section 2.2.1). By applying these approaches, L2LP moves beyond perceptual discrimination to understand the role of categories in the recognition and storage of new words in the L2 lexicon. Best and Tyler (2007) also note the different perceptual goals between discrimination and lexical recognition tasks, and the role of the lexicon in L2 perception. However, L2LP goes further, stating that L2 learning should be fundamentally modelled as meaning-driven and formalising predictions using cognitive computational models of spoken-word recognition. Additional reference is made to error-driven learning, where learning aims for optimal discrimination of cues by minimising error between cue-based expectation and desired outcome (Bovolenta & Marsden, 2022; Hoppe et al., 2022; McClelland & Rumelhart, 1981; Rescorla & Wagner, 1972; Sutton & Barto, 1998).[9]

Taken together, these three models all support the fact that the /f-v/ contrast, or at least the sound /v/, will be more difficult to acquire for L1 Arabic learners of English, based on challenges associated with new category formation over a preference for assimilation to

[9] Computational models related to natural language processing have also been used to demonstrate the need for both precision and recall in phonological representations and equivalence classification. This refers to the ability to both rapidly and accurately accept all correct exemplars of a category, as well as the precision to reject poor category exemplars (Pierrehumbert, 2016).

existing L1 categories. Despite these similarities, each model differs in their theoretical assumptions, explanatory accounts and methodological aims. Both SLM-r and L2LP argue that phonological categories are stored in abstract mental representations of sounds, while PAM-L2 "posits that perceivers extract invariants about articulatory gestures from the speech signal"(Best & Tyler, 2007, p. 20). Then, SLM-r and PAM-L2 propose a common phonological space for L1 and L2 categories, while L2LP proposes separate perception grammars for each language. PAM-L2 and L2LP approach the topic by examining L2 phonological contrasts, while SLM-r focuses on individual L2 segments. SLM-r then applies to perception and production, PAM-L2 relates specifically to perception, and L2LP extends these ideas to the developmental trajectory of lexical recognition. Finally, these models have predominantly focused their attention on naturalistic language learning, each drawing on different populations, namely: naïve or beginners (PAM), experienced or across the developmental trajectory (SLM), or simulated and human learners at varying points of development (L2LP).

The focus on naturalistic language offers some explanation for the lack of any formalised role of orthography in these models, as simultaneous exposure to spoken and written language is less of a central concern. However, the relevance of considering written input is demonstrated by evidence of orthographic influence in naturalistic learning settings (Young-Scholten & Langer, 2015) and retroactive modification of sounds initially learned without seeing orthographic input (Stoehr & Martin, 2022). A number of researchers have drawn attention to the lack of reference to orthographic influence within these models, in spite of the growing body of evidence (Bassetti, 2017; Bassetti et al., 2018; Mok et al., 2018; Nimz & Khattab, 2020; Rafat & Stevenson, 2018). In particular, Mok et al. (2018) points out that these models hinge on the perceived phonetic similarity between known and target languages, yet there is limited reference to the evidence that orthography can influence this perceived similarity. Nimz and Khattab also argue that orthography "plays a crucial role, at least when it comes to phonological representations at the lexical level" (2020, p. 23). The connections between phonological and lexical representations are particularly emphasised in L2LP; however, the PAM-L2 and SLM-r also posit lexically-driven accounts of perception and production. Therefore, if the influence of orthography is particularly critical in relation to the lexical encoding of L2 phonological categories, this critique is levelled at all three models. The next section turns to research investigating orthographic influence on L2 phonology, further demonstrating why further theoretical work is required in this area.

Before moving to the next section, it is noteworthy that Tyler (2019) has speculatively explored PAM-L2 in relation to instructed environments, including mention of written input. Tyler notes that the availability of written materials and the support for rapid L2 vocabulary expansion may reduce the opportunity for "perceptual learning", particularly in relation to difficult SC assimilations. He goes on to summarise the predicted influence of orthography as follows:

> For contrasts where learners can perceive a phonetic difference between the L2 phonemes, it is conceivable that alphabets might help learners to focus on and tune into those phonetic differences in speech, as long as the L2 orthography signals a clear phonological difference. However, in cases where the orthography does not signal a clear phonological difference, their internal rehearsal of the pronunciation of L2 words via orthography may reinforce a perception that the L2 phonemes are equivalent rather than distinct. (Tyler, 2019, p. 617)

Tyler goes on to consider situations where learners have learned a substantial amount of vocabulary via orthography, for example reading in the absence of spoken input. Increasing L2 vocabulary is then connected with fossilisation that may inhibit the improvement of L2 phonological categories, more so than would be generally predicted by PAM-L2. Overall, Tyler states that within an instructed setting, there is likely no change to TC assimilations. However, CG assimilations are less likely to be acquired, particularly if a learner rehearses words based on orthographic forms, and distinctions are likely to be restricted by rapid vocabulary acquisition. Finally, he makes it clear that if SC contrasts are difficult to acquire in immersive settings, that difficulty will be exacerbated in the classroom. While this offers a starting point for understanding the application of these models to instructed settings and the influence of written input, formalised accounts and empirical testing is required.

2.2.3 Orthographic input and L2 phonological acquisition

Despite the well-established evidence of a tight and complex relationship between literacy, phonological awareness and language processing already presented, it is only in the last few decades that SLA researchers have investigated the influence of orthography on L2 phonology in earnest. This growing body of research pays particular attention to the stark contrast between the order and amount of written language input typical of early adult instructed SLA, in comparison to child language acquisition. The prevalence of exposure to written and spoken language from the earliest stages of acquisition has led researchers to question whether this could shed light on the difficulties that adults face with L2 listening and pronunciation. Studies in this field seek to address the question of whether exposure to written forms helps or hinders phonological acquisition, falling into broad categories of positive (Erdener & Burnham, 2005; Escudero et al., 2008) and negative influence (Bassetti & Atkinson, 2015; Nimz & Khattab, 2020; Young-Scholten et al., 1999). More recent findings move beyond the binary, and have highlighted the specific aspects and contexts of phonological acquisition that are variably influenced by orthographic input (Bürki et al., 2019; Escudero et al., 2014).

Drawing together recent trends in the field, Hayes-Harb and Barrios (2021) argue four key variables mediate the influence of orthographic input (OI) on phonological acquisition, namely systematicity, familiarity, congruence, and perceptibility. These are defined as:

- **Systematicity**: whether a novel phonological contrast is systematically represented by the L2 writing system.
- **Familiarity**: Whether some or all the L2 graphemes are familiar to learners from the L1.
- **Congruence**: For familiar graphemes, whether the L1 and the L2 employ the same grapheme-phoneme correspondences.
- **Perceptibility**: The ability of learners to perceive a novel auditory contrast.

These variables clarify terminological overlaps across the field, such as the varied ways of referring to orthographic depth. Across the OI and phonology literature, the terms transparent-opaque, shallow-deep, and consistent-inconsistent are often used interchangeably. Hayes-Harb and Barrios offer “systematicity” as an all-encompassing term, which reaches beyond alphabetic representation, to focus more broadly on the written representation of L2 phonological contrasts. Here, it is important to discern whether “learners may rely on orthographic input to make target-like inferences about the phonological structure of words'' (Hayes-Harb & Barrios, 2021, p. 24). This reliability, or lack thereof, can then support or interfere with the acquisition of target phonological contrasts.

The next two variables, familiarity and congruence, introduce the relationship between the L1 and L2 orthographies. It is logical that, even if systematic written input is available, it will not support learning if it is in an unfamiliar script. Hayes-Harb and Barrios (2021) specify that graphemes are understood to be familiar to learners from the L1, which is useful to anchor a complex and dynamic concept. However, this term struggles to account for multilingualism/multiliteracy and varying degrees of familiarity. Is script familiarity based on ambient exposure, ability to decode sounds, more advanced comprehension of texts, or ability to produce writing? This raises questions about what it means to be literate in a language and indicates the insufficiency of binary categories to understand what knowledge and experience individuals bring to language learning. The idea of familiarity is easily applied to studies where participants are naïve learners, with no experience in the target language or its writing system, e.g. English learners of Arabic, Russian, and Mandarin (Hayes-Harb & Cheng, 2016; Jackson, 2016; Showalter & Hayes-Harb, 2013, 2015). Additionally, some have looked at degree of familiarity, based on similarity between L1 and L2 graphemes in Roman and Cyrillic scripts (Mathieu, 2016; Showalter, 2018) and L1 English-speakers' L2 experience in Russian and Mandarin (Hao & Yang, 2021; Hayes-Harb & Hacking, 2015; Showalter, 2020). However, it can be difficult to tease apart script familiarity from broader L2 proficiency, and the overrepresentation of L1 English monolingual samples is noteworthy.

Assuming that learners are familiar with a script, the next complicating factor is the relationship between shared symbols and the distinct sounds of each language. Returning to the example of English and Spanish: both languages have the letters <b> and <v>. However, the Spanish phonological inventory does not contain the phoneme /v/ and realises the letter <v> as /b/. This is an example of incongruence between grapheme-phoneme correspondence (GPC) rules. Incongruent L1-L2 GPCs undermine benefits associated with systematic, familiar written input, as evidenced by cross-linguistic interference and non-target production (Bassetti & Atkinson, 2015; Bürki et al., 2019; Rafat, 2016). Studies have also drawn attention to epenthesis (Bassetti & Atkinson, 2015; Silveira, 2012; Young-Scholten et al., 1999), vowel and consonant lengthening (Bassetti, 2017; Bassetti et al., 2018; Bassetti &

Atkinson, 2015), and reduced word recognition accuracy (Escudero et al., 2014; Hayes-Harb et al., 2010), in relation to orthographic input.

The final variable to consider is the perceptual salience of the target L2 sound or contrast. Instinctively, it seems plausible that OI would provide visual analysis to aid discernment of contrasts that are difficult to acoustically perceive (Escudero et al., 2008). There is also evidence to the contrary, where orthography only enhances what learners are already able to perceive (Escudero et al., 2014). Cutler (2015) claims that, even if OI facilitates the establishment of a phonological contrast in lexical representations, there may be other difficulties depending on perceptual development. For example, stored forms may not map well to perceptual processing of the speech signal. Thus, orthographically encoded distinctions in lexical representations do not necessarily improve a learners' ability to perceptually detect the phonemic distinction within an unfolding acoustic signal. These key variables are further explored in the following subsections, which present studies that closely pertain to the present study. Specifically, discussion focuses on the influence of OI (1) during the acquisition of nonnative L2 phonological contrasts in newly learned words, and (2) where the L1 and L2 do not share the same writing system.

2.2.3.1 Orthographic influence on nonnative contrasts in novel words

The influence of orthography on phonology has been investigated from numerous angles, including perception, production, processing, and the acquisition of different levels of phonology. The present study focuses on the lexical encoding of confusable nonnative phonological contrasts when learning new words. The previous sections have made reference to the encoding of phonological information in lexical representations, referring to the storage of words in the mental lexicon, most likely in declarative memory. Vocabulary research has investigated the integration of phonological, morphological and semantic information within lexical representations and how this knowledge is used in language comprehension and production (Dóczi, 2019). Additionally, reading models presented in section 2.2.1 have explored orthographic information in relation to lexical activation, retrieval, and quality. Lexical representations differ for L1 and L2 speakers, with some using the term 'fuzziness' to capture the lack of precision in phonological form and meaning, as well as the mapping between, resulting in lexical confusions based on phonological categorisation difficulties, among other phenomena (Cook et al., 2016; Darcy et al., 2013; Gor et al., 2021)[10]. The research presented in this section offers evidence of the influence of orthography on L2 phonological representations in the mental lexicon, for better and for worse. Two earlier studies emphasise the facilitative effects of orthography, in relation to L2 lexical representations and productions (Erdener & Burnham, 2005; Escudero et al., 2008). Later studies then draw attention to the specific conditions which shape the influence of orthography on L2 phonological contrasts, with a wariness towards overreliance on OI (Bürki et al., 2019; Cerni et al., 2019; Escudero, 2015; Escudero et al., 2014; Rafat, 2016; Shepperd, 2013).

Erdener and Burnham (2005) investigated the influence of orthography on the ability to produce L2 nonwords, with an additional interest in orthographic depth. 32 L1 Turkish speakers (transparent orthography) and 32 L1 Australian-English speakers (opaque orthography) were presented with 48 Spanish nonwords (transparent orthography) and 48 Irish nonwords (opaque orthography) across four experimental conditions: auditory only, auditory-visual, auditory-orthographic, and auditory-visual-orthographic, where "visual" refers to video-taped face movements. They found a facilitative effect of orthography on spoken and written production, and while visual information was also facilitative, it was overridden by orthography when they were presented together. Importantly, the authors demonstrated that the influence of orthography differed depending on L1 orthographic depth. Turkish participants exhibited less phoneme errors with transparent Spanish OI and more with opaque Irish OI than Australian-English participants, who performed similarly with both Spanish and Irish OI. Additionally, evidence that Turkish participants were more affected by orthographic information, compared to Australian participants, suggested the transfer of L1 reading processes influenced the processing of L2 orthographic information.

Rather than looking at the influence of L1 and L2 systematicity on naïve learners of pseudowords, Escudero, Hayes-Harb and Mitterer (2008) focused on perceptibility of L2 contrasts in novel words by experienced Dutch learners of English. Building on work by

[10] Orthographic knowledge is mentioned in a recently proposed Ontogenesis Model for L2 lexical representations, drawing on the concept of 'fuzziness' as a central property (Bordag et al., 2022). While theoretical attention to relationship between phonology and orthography in lexical representations is welcomed and well-overdue, this model has been criticised for a lack of testable predictions and explanatory power (Escudero & Hayes-Harb, 2022).

Weber and Cutler (2004), the authors explored the role of orthography in the asymmetric lexical activation of confusable L2 contrasts, /ɛ/ and /æ/. 50 Dutch-English bilinguals were taught 20 bisyllabic nonwords, where the first syllable of ten words minimally differed by the confusable vowels /ɛ/ and /æ/, and the rest were control words. Nonwords were presented with a "nonobject" line drawing, in either an audio-only or audio+spelling learning condition. Results from a visual world eye-tracking word recognition task revealed that the availability of OI supported the lexical encoding of an auditorily confusable L2 contrast, where spelled forms triggered asymmetric lexical activation. In contrast, those in the audio-only condition exhibited symmetrical confusion, thus lacking a lexical contrast. Like Erdener and Burnham (2005), the authors highlighted the tight connection between phonological and orthographic representations, and demonstrated orthographic facilitation in the context of congruent OI.

Continuing to investigate congruence, Escudero, Simon and Mulak (2014) highlighted the negative effects of incongruence between orthographic systems, in relation to L1 Spanish participants learning Dutch psuedowords. Proficiency was also explored by comparing 43 learners of Dutch and 30 naïve listeners. This study built on work by Hayes-Harb, Nicol and Barker (2010), who found incongruent OI had an inhibitory effect on word learning, compared to congruent OI. Their study looked at L1 English-speakers learning English nonwords, thus the encoding of nonnative contrasts was not reflected in results. Escudero et al. (2014) extended these ideas to the acquisition of pseudowords (1) differing by perceptually easy and difficult vowel contrasts, (2) with congruent and incongruent OI, and (3) learners of varying L2 proficiency. Participants were divided into audio-only or audio+ortho groups and were taught 12 Dutch psuedowords, where half differed by one of six Dutch vowels /ɪ i ɑ a ʏ y/ and the other half were fillers. An audio-visual matching task revealed Dutch proficiency significantly predicted accuracy and audio+ortho participants were more accurate with congruent OI. Importantly, experienced learners only outperformed naïve listeners in the congruent OI condition, but not with incongruent OI. Furthermore, participants in the audio+ortho condition outperformed the audio-only groups with congruent OI but performed worse with incongruent OI.

These findings were further developed by Escudero (2015), who taught Dutch pseudowords to 73 L1 Spanish and 78 L1 Australian English-speakers, incorporating orthographic depth comparisons similar to Erdener and Burnham (2005). The English speakers had no experience with Dutch, meanwhile the Spanish participants were the same as those in Escudero et al. (2014) and varied in their Dutch proficiency. The stimuli and procedure were also the same as reported in Escudero et al. (2014). Escudero again found that experience with Dutch predicted higher accuracy when learning new words that differ by perceptually difficult sounds. There was no effect of orthography on non-minimal pairs or perceptually easy vowel contrasts, as well as no effect of language background. Instead, orthography was found to improve accuracy for the two relatively easy contrasts from the perceptually difficult pairs. Escudero interpreted this as evidence that "orthography acted as a redundant or extra cue to enhance differences that could already be perceived" (2015, p. 18). The importance of strong perceptual foundations for developing lexical representations aligns with ideas proposed by PAM-L2 (Tyler, 2019) and Cutler (2015), both of whom comment on the potential for orthography to negatively impact phonological development, particularly in the case of confusable nonnative contrasts. Cutler argued that lexically encoded distinctions that cannot be reliably perceived in the input can lead to more hindrance than help, such as increased lexical competition and processing delay. Thus, lexical encoding based on segmental distinctions visualised in orthography is only useful if a contrast can be perceptually discriminated.

Rather than looking at confusable contrasts, Shepperd (2013) investigated 28 L1 English speakers' acquisition of 24 Zulu nonwords, which differed by nonnative, but perceptible, L2 click consonants. A short AXB discrimination task revealed lower accuracy discriminating nonnative contrasts (/l^{h}-gl/ and /g! - gǁ/) compared to native contrasts (/p^{h}-b/ and /v-z/); however both were discriminated well above chance. The within-subject design presented all participants with auditory forms, accompanying images and written input for half of the native and nonnative contrast items. OI adapted the Roman alphabet to form systematic representations of the words, although the click consonants were represented with the letters <C, G, Q, X>, respectively, introducing potential incongruence based on familiar L1 GPCs. Results from an audio-visual matching task found that OI significantly improved accuracy identifying native contrast items but made little difference to nonnative contrast items, in both intermediate and delayed posttests. Learner perceptions around the supportive influence of orthography also interacted with performance, where those who found OI to be supportive performed more consistently across the time delay and with different types of input.

Bürki and colleagues (2019) moved away from word recognition and demonstrated the influence of orthography on online retrieval and production of newly learned words. 26 L1 French-speakers, with varying degrees of L2 English, were taught 20 English pseudowords, where all participants saw half of the items with OI. Additionally, half the words were spelled with <i> and half with <o>, exploring the incongruence between GPC rules in English and French for these vowel letters. In a picture naming task the next day, performance was more accurate and faster for items presented with OI compared to audio-only. However, closer inspection of production data revealed that exposure to orthographic forms resulted in less native-like pronunciation of the novel items. Therefore, while encoding of the contrast in the lexicon may be supported by OI, in terms of fewer errors and faster retrieval, there is clear evidence of orthography-induced transfer where L1 and L2 GPCs are incongruent.

Rafat (2016) also investigated the influence of orthography on production patterns, where L1 and L2 share a letter but have different GPC rules. For instance, <r> for L1 English participants learning L2 Spanish pseudowords. 20 Canadian English-speakers, with no Spanish experience, were divided into either an audio+ortho or audio-only group and completed a picture naming task. The target stimuli for this study were a six-word subset from a larger study, and while stimuli were real Spanish words, they were assigned new meanings of common picturable words to ease recall. Analysis revealed that the two groups differed significantly in production patterns, where assibilated/fricative rhotics were only produced in the audio+ortho group and approximant rhotic production was higher than in the audio-only group, as well. Meanwhile, the audio-only group produced more postalveolar sibilants. Postalveolar sibliants are acoustically similar to assibilated/fricative rhotics, thus the author speculates that audio-only productions may be perceived as native-like. The audio+ortho findings are then interpreted to reflect the salience-enhancing effect of <r>, leading to the less salient feature of "rhoticity" being included in productions. However, this disambiguating effect was only found when the more salient feature, in this case assibilation/frication, was sufficiently prominent. Otherwise, it would be overridden by transfer based on L1 GPCs e.g. <r> - /ɹ/.

One further study to mention involves learning a language that does not have a distinction that is well-established in the L1, rather than vice versa; namely, the contrast between long and short consonants in Italian that is not found in English (Cerni et al., 2019). In this study L1 Italian learners of L2 English were taught a mix of English real and pseudowords, forming 30 semi-minimal pairs. In these semi-minimal pairs, the pseudoword had the same target consonant but was spelled with a single or double letter, e.g., <finish>-<prinnish> or <finnish>-<prinish>. 48 Italian high school students were assigned to an audio-only or audio+ortho condition, which involved a learning phase, test and three tasks to assess pronunciation, rhyme judgement and spelling. The audio+ortho group, but not the audio-only group, were found to produce a longer consonant when items were spelled with a double letter. Additionally, they rejected rhymes when spelling differed by either a single or double letter. The incongruence with L1 GPCs, where double letters transparently indicate geminate consonants in Italian, shows another example of orthography-induced transfer and evidence this extends to experienced L2 learners. However, they also found more novel words were learned in the audio+ortho group, aligning with a memory advantage reported in other studies (Escudero et al., 2022).

Taken together, these studies demonstrate the contextual influence of orthography on phonological acquisition of nonnative contrasts in newly learned words. While OI can support the lexical encoding of confusable contrasts (Escudero et al., 2008), these findings were in the context of high L2 proficiency, where OI was both familiar and congruent. The relevance of proficiency connects to evidence that congruent OI is more facilitative as a confirmatory cue in the context of nonnative contrasts which can be perceived (Escudero, 2015). The salience-enhancing effect of OI was also raised in relation to L2 production (Rafat, 2016). In contrast, non-target perception and production has been shown to arise from incongruence between L1 and L2 GPCs (Escudero et al., 2014; Rafat, 2016; Shepperd, 2013). Furthermore, there is clearly the possibility for both positive and negative influences of OI to be found in combination (Bürki et al., 2019; Cerni et al., 2019). The studies presented here have focused on novel word learning in the context of OI in a familiar script, shared across both L1 and L2. The next section explores effects of OI across distinct and unfamiliar orthographic scripts.

2.2.3.2 Orthographic influence across writing systems

It stands to reason that learning novel words differing by nonnative contrasts could be supported through mappings to distinct script input, as there would not be orthography-induced transfer based on incongruence. However, it is unclear to what extent learners need to be familiar with a new script and how this then relates to the perceptibility of the contrast. Limited research exists on this topic, and biscriptal literacy in bilingualism more generally (Vaid, 2022), despite the global prevalence of language learning across writing systems and orthographic scripts. However, a collection of relevant empirical studies are presented below, with a focus on learning across the Roman and Arabic scripts. To date, studies include the Roman, Cyrillic and Hangul alphabets, Arabic abjad, Japanese Kana/Kanji and Chinese Characters. Most of these studies involve L1 English-speakers who have limited or no experience with either the spoken or written language.

Starting with research involving Arabic, a series of related studies were conducted by Showalter and Hayes-Harb (2015), Mathieu (2016), and Jackson (2016), each looking at orthographic influence on the lexical encoding of nonnative contrasts by naïve L1 English learners of Arabic pseudowords. Showalter and Hayes-Harb (2015) conducted several experiments investigating whether English-speakers benefit from OI in an entirely unfamiliar script. Participants were presented with 6 minimal pairs of Arabic pseudowords, differing by the velar-uvular contrast /k-q/, and associated with a random picture from the Bank of Standardised Stimuli (Brodeur et al., 2010). In the first experiment, 30 participants were assigned to the Arabic script or audio-only condition, and completed a word learning phase, criterion test and audio-visual matching task. The second experiment followed the same procedure, except the 8 participants were given explicit instruction about the direction of writing in Arabic. In experiment 3, a further 8 participants followed the same initial procedure, but with the Roman script input. Finally, in experiment 4, the auditory items were produced by one speaker, rather than multiple speakers, and 30 participants were assigned to a new Arabic script condition and audio-only control group. Across all the experiments, there was no significant difference between the audio-only and Arabic script groups. These findings may reflect the low perceptual salience of the contrast for L1 English-speakers, or could be related to the small sample sizes. However, they then found that the Roman script group performed significantly worse than the other OI conditions.

This likely provides further evidence for orthography-induced transfer, as <k> and <q> can both map onto /k/ in English (e.g., "king" and "queen"), encouraging spurious homophony.

Mathieu (2016) extended this line of enquiry by looking at a different consonantal contrast in Arabic, namely the uvular-pharyngeal /χ - ħ/ contrast, as well as different degrees of script unfamiliarity. In his study, L1 English-speakers, with no experience of the Arabic or Cyrillic script, were assigned to an audio-only, Arabic script, Cyrillic script, and Hybrid script word learning conditions, with 21 participants in each group. The hybrid script condition involved only the target phoneme being represented by a Cyrillic letter, while the rest of the word was spelled using the Roman alphabet. Following the same procedure as Showalter and Hayes-Harb (2015), 12 Arabic pseudowords were presented with an accompanying image and OI in a word learning phase, followed by a criterion test and audio-visual matching task. As with the earlier experiments, there was no significant difference between different OI conditions, so no evidence that unfamiliar OI was inhibitory or facilitative. However, closer analysis comparing all OI conditions to the audio-only condition revealed accuracy was worse with any OI compared to no OI. These findings demonstrated that the degree of unfamiliarity was not influential on the ability to lexically encode L2 phonological contrasts. However, there was a generally inhibitory effect of entirely unfamiliar OI presentation. Mathieu also noted the perceptual salience of the L2 contrast and how the compounded difficulty of encoding a new phonological contrast represented by an unfamiliar script can negatively impact participant accuracy.

Returning to the /k-q/ contrast and potential influence of explicit instruction, Jackson's (2016) master's thesis used the same stimuli and procedure from Showalter and Hayes-Harb (2015) to investigate whether a novel grapheme or additional diacritic would facilitate the acquisition of the L2 contrast. 52 university students were assigned to four different conditions, presenting either the novel grapheme or extra diacritic both with and without instruction. Showalter and Hayes-Harb (2013) previously reported evidence that tone mark diacritics supported L1 English speakers learning novel items differing by L2 Mandarin tones. However, Jackson's results showed an advantage of learning with a novel grapheme over a diacritic. This aligns with findings that lexical stress marks did not affect the learning of novel words by L1 English inexperienced and experienced learners of L2 Russian (Hayes-Harb & Hacking, 2015). Jackson also reported a positive influence of instruction on the diacritic but not the novel grapheme group, suggesting that participants could learn to direct their attention to the unfamiliar cue. Overall, it is noteworthy that novel grapheme accuracy was comparable to the results reported in Showalter and Hayes-Harb (2015) for Arabic OI and audio-only, thus indicating a neutral influence.

Another adaptation of this design was conducted for my master's thesis and was used to pilot ideas for the present study (Shepperd, 2018). This study sought to unpack the relevance of perceptibility by comparing L1 English-speaker acquisition of 16 Arabic pseudowords, differing by either native contrasts (/s-z/ and /m-n/) or nonnative L2 contrasts (/k-q/ and /x-ɣ/). Using a mixed methods design, 28 participants were assigned to either audio-only, English script and Arabic script conditions, and completed a similar word learning, criterion test and audio-visual matching task. The matching task was adapted to integrate a visual world eye-tracking element, as well as the addition of a perception task, production tasks and post-test questionnaire. The perception task revealed high discrimination accuracy of both native and nonnative contrasts, demonstrating that issues of perceptual salience, raised in this and other studies, likely reflected the quality of the phonological representation rather than the ability to perceptually discriminate the sounds. Overall, lexical encoding accuracy was worse with any OI compared to no OI, and this was particularly pronounced with the nonnative contrast items. Meanwhile, English congruent OI improved accuracy with the native contrast words.

Eye-movement data further implied a disadvantage for Arabic OI, through less focused visual attention. Meanwhile, in the English OI group, there was a faster and stronger shift of attention to target items (match trials) and real-time processing of phonological competitors and minimal pair items (mismatch trials). English OI also supported target-like productions for both native and nonnative contrast items. Questionnaire responses then revealed negative perceptions of Arabic script influence, including increased cognitive load. In contrast, English OI was thought to positively influence learning, by both English OI and audio-only groups, encouraging explicit analysis of sub-lexical units and drawing on previous experiences learning languages. Additionally, those in the English OI group reported more learning strategies than the other conditions, with a higher reported focus on pronunciation during word learning. These findings demonstrate a complex and varied influence of script familiarity when learning new words differing by a native and nonnative contrast, where participant perceptions do not always align with task performance.

Taken together, these studies suggest that entirely unfamiliar OI does not support the encoding of L2 contrasts during novel word learning. Also, there is further evidence that congruent OI can be facilitative, whereas incongruent OI has an inhibitory effect on lexical encoding. However, this also depends on the extent to which words differ by a phonological contrast that is well-established in the L1 phonology. All these experiments were conducted with naïve participants in a single session experiment, limiting the ecological validity and pedagogical relevance of findings. It is therefore valuable to pursue insights into learning contexts and influences related to L2 proficiency and literacy from language learners with more L2 experience, or at least the intention of learning the target language. A couple of related studies demonstrate the importance of literacy experience and language proficiency from the perspective of L1 Arabic beginner learners of L2 English acquisition of syllable structure (Al Azmi, 2019) and experienced learners visual word recognition (Ota et al., 2009).

Al Azmi (2019) found that both OI and L1 literacy influenced rates of epenthesis and deletion in the production of English consonant clusters. Beginner English classes were given to 60 L1 Arabic-speakers in Saudi Arabia, which involved learning 26 target words with different English onset and coda clusters. Participants were taught in ten 20-minute lessons over five weeks, and were divided into three groups: non-literate, Arabic literate audio-only, and Arabic literate audio+ortho. The non-literate group had a higher rate of deletion than the literate groups, and the audio-only group had a higher rate of deletion than the audio+OI group. Additionally, deletion was significantly reduced when learners were exposed to OI during a posttest. These results align with research demonstrating that, while young children favour deletion over epenthesis when producing difficult consonant clusters, OI promotes epenthesis and the importance of producing all visually represented units (Bassetti, 2007; Bassetti & Atkinson, 2015; Young-Scholten et al., 1999). Importantly, these findings demonstrate that (lack of) L1 literacy experience, as well as exposure to orthography, affects the acquisition of L2 phonology.

Looking at the relationship between phonology and orthography from a different angle, Ota and colleagues (2009) used a semantic-relatedness task to demonstrate the effects of L1 phonology on L2 lexical representations during visual word recognition. They tested

L1 Arabic, Japanese and English-speakers to see whether English near-homophones would be perceived as homophonous when differing by a nonnative contrast. For example, Japanese lacks the /l-r/ distinction between <lock> and <rock>, which is likely to influence the judgement that <key> and <rock> are related. This contrast is established in Arabic, which then lacks the /p-b/ contrast found in English and Japanese. 20 participants in each group, all with high proficiency in English, were asked to judge the relatedness of 20 homophone pairs, 20 /l-r/ minimal pairs, and 20 /p-b/ minimal pairs, based on the spelling without audio. As predicted, all participants were less accurate and slower judging real homophones. Furthermore, the Japanese group was less accurate and slower with /l-r/ items, while the same was true of the Arabic group with the /p-b/ items. These findings show that cross-linguistic transfer occurs at the level of phonological representation, not misperception or production, which is striking in the context of the clear visual distinction in the spelling. Therefore, when learning across writing systems, it appears the disambiguating influence of systematic orthography is of limited use when lacking robust phonological representation of the target L2 contrast.

In addition to these studies looking at learning across English and Arabic phonologies and orthographies, several studies have examined L1 English-speakers learning of Russian pseudowords, accompanied by Cyrillic script input. Cyrillic offers the opportunity to manipulate both familiarity and congruence, as some letters of the alphabet are shared with the Roman alphabet, while others are entirely different. Following the same artificial lexicon design as replicated in multiple studies above, Showalter (2018) found that naïve L1 English-speakers were able to perform with near ceiling accuracy in a matching task, where none of the words differed by nonnative contrasts. However, in the audio+OI group, accuracy was significantly worse when accompanying OI was incongruent compared to congruent and unfamiliar conditions. Thus, incongruence is a pervasive effect, regardless of perceptibility, whereas unfamiliarity of OI does not pose additional challenges when the target contrast is established in the L1. Showalter (2020) extended this study to compare naïve, beginner and experienced learners, as well as the impact of explicit instruction. Interference effects from incongruent GPCs were particularly evident with the naïve participants, which did not improve with instruction, but these interference effects reduced with increased language experience.

Finally, several studies have investigated orthographic influence in relation to learning Mandarin with Chinese characters. These studies often compare the influence of learning with logographic characters, Romanised pinyin graphemes, or zhuyin[11] phonetic symbols for Mandarin. Hayes-Harb and Cheng (2016) assigned 30 monolingual L1 English-speakers to a pinyin (familiar OI) and zhuyin (unfamiliar OI) group, who were then taught and tested on 16 Mandarin pseudowords. Stimuli presented to the pinyin group was also either congruent or incongruent with L1 GPCs. As expected, the pinyin group performed worse when OI was incongruent with L1 GPCs, while the zhuyin was unaffected. This was interpreted to show the advantage of unfamiliar graphemes in comparison to familiar incongruent GPCs. Additionally, a perception task with target native and nonnative contrasts clarified that the results reflected a conflict between orthographic and phonological information in the lexicon, not perceptual difficulty.

Hao and Yang (2021) echoed the advantage of novel graphemes, but in the context of intermediate and advanced L1 English learners of L2 Mandarin, where characters facilitated tonal encoding more than pinyin. In fact, the opposite was true for naïve participants, demonstrating that entirely unfamiliar graphemes did not support learning. Instead, proficiency with a distinct script that systematically encodes the target phonology and does not encourage L1 transfer was beneficial. Notably, Hayes-Harb and Cheng provided some training with zhuyin, while Hao and Yang did not provide this for the characters. Mok et al. (2018) also investigated the different influence of characters and pinyin when learning L2 Mandarin tones; however, their study included L1 Cantonese-speakers who were experienced learners of Mandarin. They found that pinyin facilitated the perception and production of tones in monosyllabic words, whereas characters proved more beneficial when it came to disyllabic words. Furthermore, it was demonstrated that participants with lower task performance were more affected by orthographic information than those with higher performance. Each of these studies investigating the learning of Mandarin provides further evidence that L2 proficiency is a relevant variable to consider when looking at orthographic influence, including L2 script literacy experience. Also, while they all offer some evidence of the benefits of distinct OI, the extent of this effect remains unclear without comparison to an audio-only condition.

[11] Zhuyin is typically found in Taiwan, whereas pinyin is more common in mainland China.

In summary, this section has brought together evidence that orthographic influence on L2 phonology persists across writing systems, specifically where the L1 and L2 are represented by distinct scripts. While some studies suggest that distinct/unfamiliar OI is beneficial over shared incongruent OI (Hayes-Harb & Cheng, 2016; Jackson, 2016; Showalter, 2018), others present evidence that this influence is more neutral (Showalter & Hayes-Harb, 2015) or negative (Mathieu, 2016; Shepperd, 2013), when compared to learning without OI . As has been noted with orthographic influence more generally, the extent and nature of this influence is connected to L2 proficiency (Hao & Yang, 2021; Mok et al., 2018; Showalter, 2020), L1 literacy experience (Al Azmi, 2019) and encoding contrasts that are not established in the L1 (Mathieu, 2016; Ota et al., 2009; Shepperd, 2013). The limited research and mixed findings presented in this section demonstrate the need for further investigation into this complex area, particularly with larger sample sizes and including learners from more diverse backgrounds than L1 English-speaking undergraduate populations.

2.2.4 Summary of orthographic influence on phonological development

This section delved deeper into theoretical accounts and empirical research that pertains to the relationship between orthography and phonology, in the context of language learning. As has been stated throughout this review, one of the defining features of SLA is that at least one other language has already been acquired. Therefore, L2 learning never involves a blank slate, but rather makes constant reference to existing linguistic knowledge and is contextualised within broader cognitive developments. For example, language specific phonological attunement occurs early in infant language acquisition. Literacy development also commonly begins during childhood, as well as memory and learning systems changing over the lifespan. L1 literacy then influences phonological development in two critical ways. Firstly, literacy experience is directly related to our metalinguistic, particularly phonological, awareness. This then connects to our understanding of how continuous speech can be segmented and represented in smaller units. Secondly, the entrenched associations between specific written and auditory forms can promote crosslinguistic interference, which is particularly evident when grapheme-phoneme correspondences are incongruent across known and target languages.

Despite the clear connection between literacy experience and phonological development, in both L1 and L2 contexts (section 2.2.1), formalisation of orthographic influence is notably absent from speech learning theories (section 2.2.2). Growing empirical evidence focusing on orthographic influence and L2 phonology make a strong case for the need to develop both explanatory and predictive accounts to direct future research (section 2.2.3). That being said, there is already much scope to expand the existing literature and clarify contradictory and inconclusive findings. In particular, language learning across distinct rather than shared writing systems deserves further exploration. This is a core motivation for the present study, which is more fully outlined in section 2.4, alongside the guiding research questions for this thesis. The next section details the phonological and orthographic systems of the languages included in the current study.

2.3 Arabic and English language and literacy

As the present study focuses on the acquisition of English consonant contrasts by Arabic-speakers, in relation to Arabic and English script input, phonological and orthographic representations of speech segments in both languages are outlined below. Additionally, broad observations about the two languages are made, with specific reference to the relationship between speech and writing, to provide further context for ideas developed throughout the thesis.

2.3.1 Arabic phonology and orthography

A fundamental characteristic of Arabic is that it is a diglossic language. This refers to "a situation with the coexistence of two related varieties considered to belong to the same language, with one variety having exclusively formal uses" (Mejdell, 2018). Early descriptions of Arabic diglossia by French scholars in the 1930s characterised the distinction between varieties as a written and spoken Arabic. However, Ferguson (1959) influentially redefined the term with hierarchical concepts of formality and prestige attached to the standard 'high' variety, in contrast to colloquial 'low' varieties.[12] This functional distinction

[12] There is much debate around whether varieties in diglossia are of the same language, differences between diglossia and bilingualism, and whether anyone can consider Standard Arabic to be their mother tongue (Mejdell, 2018).

is commonly referred to by Arabic speakers as *al-'arabiyya al-fuṣhā* (standard) and *al-arabiyya al-'āmmiyya* (colloquial). As hinted at by the earlier classifications, 'standard' Arabic is typically learned in educational and religious settings, as well as being predominantly used in written and formal contexts. It is not used in ordinary conversation. There is also a strong link between Classical Arabic and Islam, dating back to the revelation of the *Qur'ān*, with historically embedded ideas around the importance of written texts, as well as memorisation and correct pronunciation for proper recital of its passages (Spolsky, 2003; Wahba, 2021). In contrast, colloquial varieties are primarily spoken languages and lack standardised written representations. Due to their prestige, Classical Arabic and Modern Standard Arabic (MSA) have typically been the focus of scholarly interest, both inside and outside of Arab countries, with considerably less attention being paid to regional varieties (Procházka, 2021). When referring to Arabic diglossia throughout the discussion below, the terms MSA and colloquial/regional varieties are used.

At least 300 million people speak Arabic as a native language (Owens, 2013). The Arabic-speaking world stretches from Oman to Mauritania, reaching northwards to Turkey and then south to Chad. Beyond the boundaries of Arab nation states, Arabic-speaking minorities are found all around the world, particularly in Central Africa and Central Asia. In addition to understanding the relationship between MSA and regional varieties, it is important to consider how colloquial dialects relate to each other. For example, "though speakers of adjacent dialects can usually understand each other, dialects situated far apart may be mutually unintelligible"(Procházka, 2021, p. 220). Colloquial varieties can broadly be divided into western 'maghreb' and eastern 'mashreq' dialects, where the former refers to North Africa and the later to the Middle East/West Asia. Figure 2.5 demonstrates the geographical spread and further classification of Arabic dialects with shared linguistic features into: Maghreb (North Africa), Egypt-Sudan, Levantine (Jordan, Syria, Lebanon, Palestine), Mesopotamia (Iraq), and Arabian Peninsula/Gulf (including Saudi Arabia, Oman, Qatar etc.) dialects (G. Brown & Hellmuth, 2022)[13]. These classifications reflect divisions between nomadic and sedentary populations, as well as geography. For example, nomadic dialects tend to more closely resemble MSA, while sedentary dialects show more influence from languages which have existed alongside Arabic in major cities (Mustafawi, 2017).

[13] As well as between dialect variation, there is also variation within each dialect (Mustafawi, 2017)

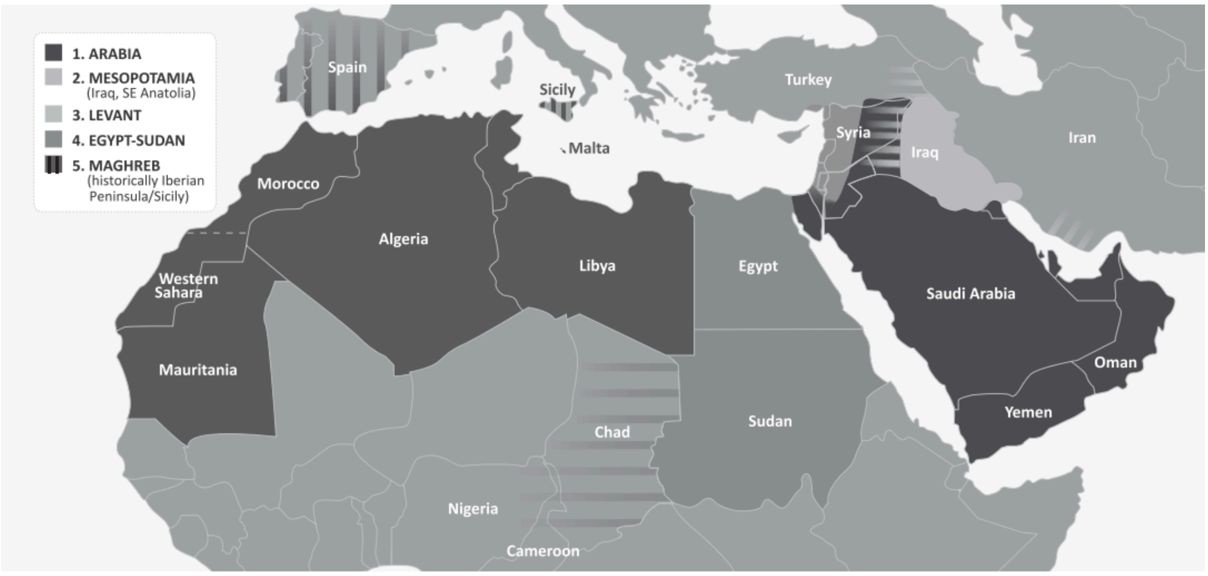


Figure 2.5: Map of Arabic dialects from Procházka (2021, p. 222)

All varieties of Arabic have large consonant inventories and a limited number of vowels, as is common with semitic languages (Mustafawi, 2017). The MSA phonemic inventory consists of 28 consonants and six vowels. The consonant inventory for MSA is displayed comparatively alongside British English consonants in table 2.2. A notable feature of Arabic phonology is the large proportion of 'guttural' consonants, including post-velar /q χ ʁ ħ ʕ h ʔ/ and emphatic /ṭ ḍ ṣ ẓ / consonants, which involve post-velar secondary articulation (Hellmuth, 2013). Comparison of these symbols with those in table 2.2, as well as the inventory provided by Mustafawi (2017), highlights points of discussion in the Arabic phonology literature around representing pharyngealization of emphatics (e.g. /tˤ dˤ sˤ ðˤ/) and velar (e.g. /x ɣ/) or post-velar fricatives. Common differences between MSA and regional dialects include sound changes with interdental fricatives (e.g. /θ/→ /t/ or /s/; /ð/ → /d/ or /z/), voiced palatal-alveolar affricate /dʒ/ realisation as /g/ in Egyptian Arabic and /ʒ/ in Gulf Arabic, and varied realisation of /q/ as /ʔ/ or /g/, depending on region and urban-rural distinctions. In relation to the present study and the target sounds highlighted in table 2.2, the /m-n/ nasal contrast is reliably found across all dialects (Smith, 2001; Watson, 2002), whereas /f/ but not /v/ is typically found in Arabic varieties and presents well-documented confusion for Arabic learners of English (Altaha, 1995; Binturki, 2008; do Val Barros, 2003; Hassan, 2014; Kharma & Hajjaj, 1989; Smith, 2001; Tushyeh, 1996). However, the colonial influence of French in North Africa means that the /f-v/ contrast is more familiar in Magrebi and Egyptian varieties, through loanwords, bilingual education, and general language contact. It is also relevant to note that the Arabic lexicon is typically understood in terms of three-letter consonantal roots, which then take on different affixation and vowel patterns to form individual words (Hellmuth, 2013)[14]. An example of roots and patterns in Arabic is given in figure 2.6.

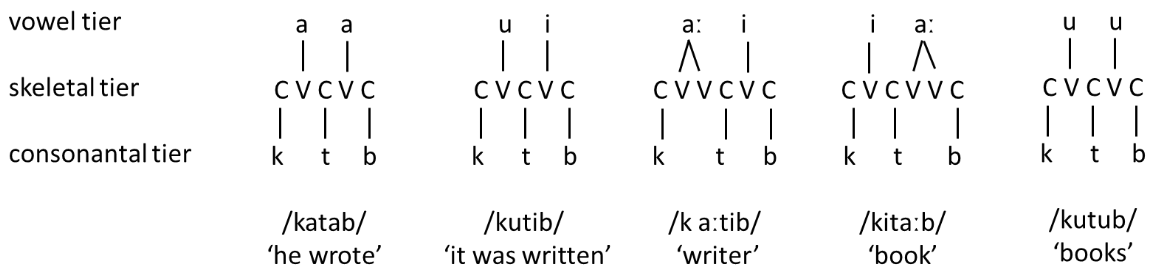


Figure 2.6: Example of words deriving from the root /k-t-b/, including semantic connection and patterns of vowels for each lexical item

[14] There are additional restrictions within consonantal roots regarding the co-occurrence of identical consonants in the first two consonants, but not the second and third (e.g., *<mmd> vs. <mdd>), as well as varying restrictions around the co-occurrence of consonants which share their place of articulation.

Table 2.2: Overview of MSA and British English consonant inventories

Manner of articulation	Place of articulation									
	Bilabial	Labio-dental	Dental	Alveolar	Post-alveolar	Palatal	Velar	Uvular	Pharyngeal	Glottal
plosive	**b**[1] p b[2]			**t d** t d **tˤ dˤ**			**k** k g	**q**		**ʔ**
affricate					**dʒ** tʃ dʒ					
nasal	**m**[3] m			**n** n			ŋ			
trill				**r**						
fricative		**f** f v	**θ ð** θ ð **ðˤ**	**s z** s z **sˤ**	**ʃ ʒ** ʃ ʒ		**x ɣ**		**ħ ʕ**	**h** h
approximant					ɹ	**j** j	**w** w			
lateral				**l** l						

[1] MSA consonants on the left in bold. Adapted from Thelwell and Sa'adeddin (1990). Analysis was based on a speaker from Palestine, who lived and studied in Lebanon and Syria; thus, likely to reflect elements of Levantine Arabic.
[2] British English consonants on the right. Adapted from Roach (2004).
[3] Target contrasts for the present study outlined in red.

Research investigating Arabic phonological development in children is increasingly turning its attention to the interplay between the diglossic varieties and how that impacts language acquisition. For example, the consonantal inventory of a child learning Arabic expands with the acquisition of MSA and, in contrast to colloquial Arabic, it is predominantly learned through reading and writing (Froud & Khamis-Dakwar, 2021). Ambient exposure to MSA is present from television, religious contexts, and reading from books from an early stage, with research demonstrating usage of MSA in preschool children (Albirini, 2018). However, as has been shown, L1 literacy development is tightly connected to phonological development, particularly phonological awareness. For example, Saiegh-Haddad (2004) has shown that children in the first two years of school take longer to learn letters that do not correspond to sounds in their colloquial variety. Furthermore, Eviatar and Ibrahim (2014) responded to evidence that skilled reading and reading acquisition is slower in Arabic compared to other languages by suggesting that diglossia and the complexity of Arabic orthography are contributing factors.

Arabic is written in the Arabic script, which is also used to represent Persian languages, Urdu, Pashto, and Kurdish, among other languages. Reference to Arabic spelling and script in this study specifically relates to Arabic orthography. Notable features of the orthography include the direction of writing from right-to-left and the writing of verbs and nouns reflecting consonantal roots and patterns. Short vowels are mostly absent from written words, except in children’s books, educational materials, and religious texts, where diacritic marks are added to fully specify the phonological form of the word. This has led scholars to describe the orthographic depth of Arabic as shallow when it is vowelled, and deep when a text is unvowelled, as the full phonological form must be inferred from context and lexical representations (Eviatar & Ibrahim, 2014). This is why Arabic is better described as an abjad or consonantal alphabet. For example, taking the words from figure 2.6, ’he wrote’ <كَتَب >, ‘it was written’ <كُتِب>, and ‘books’ <كُتُب> would all have the same unvowelled spelling i.e., <كتب>. In contrast, the long vowels in ‘book’ <كتاب> and ‘writer’ <كاتب> would always be represented in the spelling.

In addition to the absence of short vowels, the letter forms themselves are complex, in that the location and number of dots is integral to many letters, as well as many letters having different shapes, depending on their position in a word (see table 2.3). Moreover, while there is a transparent one-to-one correspondence between graphemes and phonemes, similar graphemes rarely represent similar phonemes (e.g., < ر > -/r/ and <ز> - /z/) and the same phoneme can be represented by multiple letter forms (e.g., /h/ - < ه ـه ـهـ هـ >). Further, most letters are connected to both right and left letters in a word, but six letters are only connected to a preceding but not following letter. Of course, Arabic-speakers develop fluent literacy and reading strategies appropriate for the orthographic system. Also, as noted in section 2.2.1, different writing systems bring varying benefits and challenges to reading processing, where one-to-one alphabetic representation is not always the best fit. Indeed, an abjad is well-suited for efficient representation of the Arabic root and pattern systems.

Table 2.3: Overview of Arabic graphemes, letter forms and phoneme correspondence

Grapheme	Connected letter form	Phoneme
ا	ـا	a
ب	ببب	b
ت	تتت	t
ث	ثثث	θ
ج	ججج	d͡ʒ
ح	ححح	ħ
خ	خخخ	x
د	ـد	d
ذ	ـذ	ð
ر	ـر	r
ز	ـز	z
س	سسس	s
ش	ششش	ʃ
ص	صصص	sˤ
ض	ضضض	dˤ
ط	ططط	tˤ
ظ	ظظظ	ðˤ
ع	ععع	ʕ
غ	غغغ	ɣ
ف	ففف	f
ق	ققق	q
ك	ككك	k
ل	للل	l

م	ممم	m
ن	ننن	n
ه	ههه	h
و	و	w
ي	ييي	j
ء	ء	ʔ

2.3.2 English phonology and orthography

While there are many varieties of English found around the world, arising from differing sociolinguistic contexts and with varying linguistic structures (Kachru, 1997), English is not a diglossic language. That being said, the focus of this overview will be on British English, specifically Southern Standard British English (SSBE), as this was the variety presented in the majority of materials for the current study. Like Arabic, English also has a large consonantal inventory, with around 23 consonants. However, there are less guttural consonants compared to Arabic, among other differences (table 2.2). English has a large vowel inventory as well, with around 20 vowel sounds in SSBE. Unsurprisingly, English sounds that do not appear in the Arabic phonological inventory, particularly vowels, prove difficult for Arabic learners of English to identify (Evans & Alshangiti, 2018; Shafiro et al., 2012). As the present study is primarily concerned with consonants, vowels are not discussed here, apart from to highlight the opacity of written representation of vowels in English (e.g., four times as many vowel sounds as letters). For Arabic-speaking learners of English, this is then complicated by the transfer of reading strategies that give lower priority to written vowel information (Hayes-Harb, 2006).

English is represented by the Roman/Latin alphabet, alongside numerous languages around the world. As the present study is concerned particularly with English orthography, the terms English spelling/script are used, rather than the Roman alphabet. When discussing literacy development and orthographic depth in section 2.2.1, English was highlighted as having a notoriously deep orthography, where the one-to-many mappings are exemplified in table 2.4. The table also demonstrates that those acquiring literacy in English need to learn both uppercase and lowercase forms, as well as the use of silent letters, double letters, and digraphs to represent single phonemes. Therefore, while Arabic is characterised as a deep orthography by some, its manifestation of orthographic depth is distinct from English orthography. In Arabic, the opacity of orthography is in absence, drawing on inferential reading strategies. Meanwhile, English grapheme-phoneme correspondences are unreliable and often misleading if a one-to-one correspondence is assumed. It is noteworthy that Romance languages, such as Italian and Spanish, are more transparently represented by the Roman alphabet. However, the application of these same letters to Old English did not provide the same fit to the phonemes of a Germanic language. English spelling does not systematically map to the sounds of the language, but rather reflects a history of language contact and the surrounding context.

Table 2.4: Example grapheme-phoneme correspondences and letter forms for English consonants, including examples with varying word positioning

Phonemes	Graphemes	Examples
p	P p	**p**ot, sto**pp**ing, stee**p**
b	B b	**b**ig, bu**bb**le, so**b**
t	T t	**t**ea, se**tt**le, spli**t**
d	D d	**d**og, ri**dd**le, roun**d**
k	C c/ K k/ -ck/ -ch/ Q q	**c**oot, su**cc**ess, **kick,** ban**k,** a**ch**e, **q**ueen
g	G g	**g**o, so**gg**y, bo**g**
d͡ʒ	J j / G g/ -d(i/u)/ -dg	**j**u**dg**e, **g**in, ma**g**ic, sol**di**er, e**d**ucate
t͡ʃ	Ch ch/ C c/ -t	**ch**ur**ch**, **c**ello, na**t**ure
m	M m	**m**oon, nu**m**ber, swi**mm**ing
n	N n/ Kn kn	**n**ote, ru**nn**ing, **kn**ow**n**
ŋ	-ng/ -n(k)	Si**ng**, li**ng**er, dri**n**k
f	F f/ Ph ph/ -gh	**f**lu**ff**, **ph**oto, dol**ph**in, enou**gh**
v	V v/ -f	**v**al**v**e, o**f**
θ	Th th	**th**rough, mo**th**
ð	Th th	**th**ough, mo**th**er
s	S s/ C c	**s**ock, gla**ss**, **c**ity, a**c**e
z	Z z/ X x/ -s	**z**oo, da**z**e, ja**zz**, **x**ylophone, pro**s**e
ʃ	Sh sh/ S s/ -ti	**sh**op, fi**sh**, **s**ure, emo**ti**on
ʒ	-s/ -z/ -ge/ -ti	mea**s**ure, vi**s**ion, sei**z**ure, bei**g**e, equa**ti**on
h	H h/ Wh wh	**h**and, re**h**ab, **wh**ole
ɹ	R r	**r**ed, a**r**oma, t**r**ue, qua**rr**y
l	L l	**l**ittle, fu**ll**
j	Y y	**y**ellow, **y**o**y**o
w	W w/ Wh wh	**w**aste, **wh**at, allo**w**
ks	X x	a**x**e, bo**x**

A better way to understand English spelling is to consider it polysystemic, combining the spelling systems of Anglo-Saxon, Scandinavian, French, Latin and Greek, through the vehicle of the Roman alphabet (Upward & Davidson, 2011). Additionally, the borrowings of spelling from an even wider range of languages can be seen in loanwords, such as tsunami, khaki, myrrh, czar, fjord etc. Others have applied the term morphophonemic to English spelling, as representations best reflect a combination of morphology, etymology, and phonology (Bowers & Bowers, 2017). As Venezky puts it:

> English orthography is not a failed phonetic transcription system, invented out of madness or perversity. Instead, it is a more complex system that preserves bits of history (i.e. etymology), facilitates understanding, and also translates into sounds. (Venezky, 1999, p. 4)

For example, affixation of English words can be flexibly combined to encode meaning and function, as well as indicating the history of the word. Take the word *sign*, meaning 'mark/gesture/token'. Affixes can be added providing reliable morphological cues, e.g., *de*sign, *rede*sign, *de*sign*ed*, sign*ature*, sign*ify*, sign*ificant* etc., which in turn provide more consistent phonological mappings when this is considered. For this reason, more attention to visible morphology in spelling has been advocated for, in order to understand skilled reading in English (Rastle, 2019). The issue of orthographic transparency and morphophonemic representation is of tangential significance in the present study, as all pseudowords are presented with reliable one-to-one GPCs of target contrasts, to focus on perceptual salience and familiarity in the context of English orthographic input.

The information presented here in relation to Arabic and English sound and writing systems is in no way comprehensive, but serves to briefly introduce contextual insight for the present study. The crucial points to bear in mind when considering the research aims and questions outlined in the next section are that Arabic and English differ in their consonantal inventories and their orthographic systems. Arabic varieties typically lack the voicing contrast between interdental fricatives /f-v/ and Arabic-speaking learners of English struggle to perceive and produce this contrast. Additionally, Arabic uses an abjad that varies in its orthographic depth, depending on use of vowel diacritics, and English uses an alphabet which is opaque in its GPCs, encoding additional morphological and etymological patterns. Thus, L1 Arabic learners of L2 English are navigating word learning involving distinct phonological contrasts, across different scripts, with varying consistency of representation, and divergent habituated reading strategies. The present study seeks to shed light on the complex relationship between the written and spoken forms of these languages in the minds of language learners.

2.4 Rationale and research questions

A clear case has been made for the importance of considering input when investigating second language learning, offering insight into acquisition, processing, and representation of additional languages throughout the lifespan. Child and adult language learning differs in several important ways. Some variables include the timing, quality, and type of input, as well as noticeable differences in ultimate attainment. Specifically, modality of input (written vs. spoken) and persistent difficulty perceiving, processing and producing nonnative speech sounds have been highlighted. Adult language learning, in instructed settings, is often characterised by early exposure to large quantities of written input alongside, or even exceeding, exposure to spoken language. The reverse is true during child language learning, as spoken language has primacy and literacy commonly develops after the phonology of infant language is well established. The importance of making this comparison is that literacy and phonology are tightly related, in both L1 and L2 development. This logically connects to the fact that the written representation of a language affects the way we think about, remember, and process the sounds of that language.

The present study seeks to investigate the relationship between exposure to written input and the challenges that adult language learners face regarding target-like phonological development, particularly the learning of new words that differ by confusable nonnative phonological contrasts. Research investigating the influence of orthographic input on L2 phonology has expanded greatly in the last twenty years, with consistent evidence that systematicity, familiarity, congruence and perceptibility are important contributing variables to orthographic influence. One of the most robust findings is that incongruence between the GPCs of known and target languages, which share the same orthographic script, regularly results in orthography-induced L1 transfer and interference, which persists into high levels of proficiency. As this is an emerging field, findings relating to familiarity, perceptibility, amongst other points, are less conclusive. Furthermore, limitations of the field include a focus on languages that share the Roman alphabet, overrepresentation of monolingual L1 English-speaking undergraduate participants, and small sample sizes.

In order to build on existing research and address these frequent limitations, the present study investigates the influence of orthographic input when the L1 and L2 do not share the same orthographic script. Additionally, it includes learners often underrepresented in language research, namely L1 Arabic-speakers from a range of national and educational backgrounds. Participants also varied in their English language and literacy experience, offering insight into proficiency and biscriptal language learning. Extending the work of related studies, a psycholinguistic design involving novel word learning and an audio-visual matching task was adapted for use with internet-based methods. While there are many benefits to using internet-based methods, one important reason was to investigate this topic with a larger and more diverse sample.

The evidence presented in this review of the literature also highlighted the need to consider the awareness of and attention to input, as well as the language experience that adult learners bring with them to SLA. Part of this experience usually includes literacy, which is related to phonological awareness, reading strategies, and broader beliefs about the significance of written language for learning. To my knowledge, there is no existing research systematically investigating participants' perceptions about their learning with written input and how that relates to L2 phonological acquisition during word learning. The present study thus proposes a novel direction of inquiry in the field, by integrating participant perspectives on orthographic influence during early word learning and their language learning strategies. This exploratory angle and combination of deductive and inductive approaches is advantageous for gaining a better understanding of the topic, as well as moving towards a theoretical account for an often-overlooked area. Finally, the insights from this study aim to better understand the influence of written input in a way that can be applied to instructed L2 settings, particularly when teachers and learners are navigating diverse language and education backgrounds across writing systems. The research questions directing the present study are as follows:

Primary research question: Does exposure to the Arabic and English script differentially influence L1 Arabic-speakers' acquisition of new words that differ by a difficult nonnative phonological contrast?

Secondary research questions:

1) To what extent do different types of orthographic input (OI) influence the lexical encoding of L2 phonological contrasts in memory during novel word learning?
2) What influence of OI were learners aware of during word learning?
    a) What was the perceived influence of OI during the present study?
    b) What were the participants' beliefs about the importance of OI when learning new words more broadly?
3) What is the relationship between learners' language learning strategies and lexical encoding of L2 phonological contrasts in novel words?

Based on the evidence presented in this chapter, it is anticipated that exposure to written input using the L1 Arabic script to spell novel words will have a different effect during word learning and recognition compared to the L2 English script input. Specifically, when learning a difficult nonnative phonological contrast, such as /f-v/, the mapping of the two sounds to a single <ف> letter in Arabic will result in poor accuracy lexically encoding the contrast, compared to the transparent mapping to two separate <f> and <v> letters in English. However, this is likely to vary depending on how familiar L2 learners are with both the phonology and orthography of English. Additionally, it is expected that participants will report a preference for exposure to written input over audio-only input alone and will place a high value on visual representations when learning new words in a language. Finally, it is predicted that participants will actively make use of a range of language learning strategies, drawing specifically on orthographic information, amongst other strategies. The reasons behind participant perceptions and the precise strategies that they make use of remains to be seen, as does the relationship between perceptions, strategies, and performance in the study. The next chapter outlines the methodology, before then presenting the analysis and findings for each research question in turn.

Chapter 3: Methodology

3.1 Approach

The present study employs a within-subjects, mixed-methods design that seeks to rigorously explore the influence of written input on adult phonological development during L2 word learning, by combining psycholinguistic techniques with qualitative insights into participants perceptions and strategies. The present design draws inspiration from studies investigating L2 orthography and phonology (Escudero, 2015; Escudero et al., 2008; Rafat & Stevenson, 2018), including research that focuses specifically on unfamiliar written input (Hayes-Harb & Cheng, 2016; Hayes-Harb & Hacking, 2015; Mathieu, 2016; Showalter & Hayes-Harb, 2013). The combination of measures integrated in the experimental instrument aspires to disentangle the complex influences of multimodal input processing during novel word learning.

The present study argues that inquiry into participant reflections is necessary to better understand learners' conscious awareness of written input, its perceived influence on learning, and the strategies applied to learning which make use of the phonological and orthographic information available from the input. As noted in chapter 2, the majority of related research does not include participant perspectives and recruits from pools of university students, with an overrepresentation of L1 English-speaking participants. This raises serious concerns for representation and generalisability of research (Andringa & Godfroid, 2020; Plonsky, 2023; Shepperd, 2022). In contrast, the present study embraces the opportunity to cast the net wider through internet-based[15] data collection.

In response to a general underarticulation of ontological and epistemological assumptions in the field of psycholinguistics, which has a historical bias towards positivism, behaviourism and cognitivism (O'Connell & Kowal, 2003), this project incorporates views that have emerged from the 'social turn' in second language acquisition research (Block, 2003; Ortega, 2011). Within these debates, there has been a call to 'bridge the gap' between the cognitive and the social, especially with regards to epistemological differences often

[15] The terms *internet-based* and *online* are used interchangeably throughout the thesis. However, *internet-based* is preferred and particularly used in contexts where there could be confusion with the psycholinguistic concept of *online*, in reference to real-time processing.

posited between quantitative and qualitative researchers, which are often equated with cognitive and sociocultural perspectives respectively (Hulstijn et al., 2014). It takes little probing to reveal the false equivalence between these approaches, as quantitative methods are clearly found with both cognitive and sociocultural research, and increasingly the incorporation of qualitative methods and social perspectives are becoming commonplace within cognitive research. Furthermore, there are clear grounds to argue that there is in fact no such gap in SLA to be bridged, as is highlighted by the following:

> Obviously, most language learning takes place in social and cultural contexts that play a role in the learning process (and all language learning does, if you call the psycholinguistics laboratory a sociocultural context, too, albeit a highly marked one). Equally obviously, all social and cultural learning is at least in part a cognitive process: we acquire social and cultural knowledge, and knowledge and cognition are the same thing. (Hulstijn et al., 2014, p. 365)

While I do not agree that knowledge and cognition are obviously synonymous, cognition is near impossible to separate from sociocultural context. The perceived gap is more indicative of intradisciplinary boundary lines within the fields of psychology and linguistics, as well as the struggle to effectively communicate between theory and practice.

To this end, a critical realist approach (Bhaskar, 2009; Price & Martin, 2018) is adopted, with the underlying ontological assumptions that processes involved in language learning, and specifically phonological development and word learning, are real and exist independent of the researcher's theories and constructions. These processes are assumed to be socially influenced cognitive processes, where experiences of the sound and writing systems in the first and second language interact with oral and written language exposure. To better understand the nature of these interactions, observable events generated by these processes are investigated, alongside complementary insight gained through reported perceptions and experiences. A mixed-method approach is applied to address the epistemological assumption that what we can expect to understand about underlying phonological processes is imperfect and inevitably biased by both the experimental instruments and the researcher's personal axiology. The latter point motivates reflexivity in relation to methodological, analytical, and interpretational choices, reflected in the positionality statement below (3.1.1).

The observable events in question are the ability to perceptually discriminate target sounds, establish and retain categorical distinctions in memory and accurately detect the distinction between novel lexical items which differ by the target contrast. This behavioural data is then complemented by self-reported insights from the participants' perceptions, experiences, and strategies. While it is important to acknowledge that an experimental setting is far-removed from both naturalistic language learning and the language classroom, the rigour and reflexivity applied throughout the study demonstrate the strength of this mixed methods design. A final note regarding Critical Realism is the connection with emancipatory critique (Bhaskar, 2009; Price & Martin, 2018), which proposes that social science contributes to broader debates, not only describing and explaining the way things are but also the moral imperative to offer informed improvements and alternatives. In the context of the present study, this relates to the assumption that the reliance on literacy in instructional settings, particularly during early language learning, is socio-politically or ideologically neutral and psycholinguistically beneficial. Thus, the results of this study and insights into socio-cognitive processes are discussed in terms of application with language education settings, as well as broader theoretical implications.

3.1.1 Statement of positionality

Firstly, as a native speaker of Standard Southern British English (SSBE) with over a decade of experience teaching the English language in the UK, Spain and Jordan, I recognise the hegemony of English as a global language and the prestige attached to my particular variety of English. In addition, my entire educational experience has been based in the UK, except for 6 months of intensive Arabic language study in two educational institutions in Jordan. However, the curriculum of these institutions was also directed at a Western audience. I acknowledge that my context is distinct from the participants in the present study; however, I bring my experiences as a language learner, teacher, and researcher to the design and analysis. I bring insights from my personal experiences learning Arabic and teaching Arabic-speakers, as well as learning other (mostly European) languages and teaching speakers of many diverse linguistic backgrounds. Additionally, I bring my knowledge and expertise in the fields of cognitive, theoretical, and applied linguistics.

Part of my motivation for learning Arabic was the importance I place on research ethics and inclusivity. I felt it was important to be able to communicate directly with participants and ensure that translations were both accurate and appropriate. I also wanted participants to know I was genuinely interested in learning from them, about their language, culture, and educational experiences. I also sought to increase my sensitivity to the barriers to participation and risks for more vulnerable participants, such as those with complex migration backgrounds, histories of trauma, and limited English language and literacy skills. Here, vulnerability is not viewed as a single category or construct that can be equally applied to particular groups of people. Instead, it is understood as varying layers, which can overlap and change depending on the individual and particular context (Luna, 2009, 2019). In this case, I sought to reduce issues around informed consent, psychological harm, and scientific integrity by improving my Arabic language skills. More information about ethical considerations is provided in section 3.7. While my Arabic language proficiency proved invaluable when conducting this research and analysing findings, it is ultimately limited, as is my cultural understanding. For this reason I sought active collaboration and reflective discussion with L1 Arabic-speaking colleagues from a range of national and cultural backgrounds throughout the process.

My perspective is additionally influenced by my experiences of language learning in contexts of migration. Firstly, this was as an adult language learner and migrant in Spain and Jordan, during which time I developed beliefs about my personal language learning experiences in relation to my academic background in linguistics. Secondly, my beliefs related to the influence of written input in SLA have evolved in response to my experiences teaching English in multilingual and multi-scriptal migrant classrooms, including adults with limited L1 literacy. These experiences resulted in a sense of urgency to help equip learners and teachers with more knowledge and better tools to critically engage with assumptions around literacy and language learning. I also developed a profound scepticism of the heavy reliance on early written input; witnessing how this can disadvantage those already marginalised in the language classroom.

The present study reflects the value I place on robust and diverse samples, as has been highlighted several times. Improved representation of diverse language learners in psycholinguistic research relates to my belief that linguistic science can provide insights into learning and processing that can then benefit language teachers and learners. As sociocultural context interacts with cognition, it is important to include learners with diverse backgrounds to understand the generalisability of findings and appropriate adaptation for instructed settings. The use of internet-based methods facilitates both of these aims, as restrictions are removed around geographical proximity, physical accessibility and individual availability (see section 3.8). Extended discussion of the inclusivity of internet-based methods for underrepresented language learners in psycholinguistic research, based on the present study, can be found in Shepperd (2022). Additionally, I have encountered views that participant perspectives about their learning and processing are of limited value, as they are subjective, incomplete, and unreliable. In opposition to this view, I trust that participants can uniquely provide insight into the context of their learning and conscious processing.

Finally, within this study there is a commitment to open science, where materials, data and analysis will be made available through Gorilla and OSF, as soon as possible. The value I place on transparency and reproducibility in research is also reflected in the appendices, where I have included adaptations of R Markdown (Rmd) files with the R script and explanations required to conduct the quantitative analyses in this project. Additionally, full transcripts of coded qualitative data are included. Thus, the accompanying appendices are lengthy and not necessary for comprehension of the research presented here. However, they demonstrate the future reproducible outputs of this work.

Overall, this section has sought to demonstrate the reflexivity and rigour applied at every stage of this study and the values embedded in the research decisions reported below. The subsequent sections provide an overview of the piloting, participants, stimuli design, procedure and research instruments used, followed by ethical considerations and issues relevant to internet-based methods.

3.2 Pilot

Piloting of the present study was conducted iteratively, as a number of changes needed to be made to the original design of the project in response to Covid-19 restrictions and the move to internet-based methods. This section presents a brief overview of two pilot phases, which led to the final research design.

3.2.1 Pilot study 1

The first pilot involved a multi-session experiment using Gorilla experiment builder and was completed by three L1 Arabic-speaking participants. Participants were presented with 18 English pseudowords, differing by three contrasts (e.g., /f-v/, /m-n/ and /θ-ð/), in a within-subjects design. The /θ-ð/ contrast was initially chosen as an example of a contrast that is established in both languages but the two sounds are represented by the same digraph <th> in English spelling. Words were presented with an image and either Arabic spelling, English spelling, or no spelling. The first three sessions involved a similar word learning and audio-visual matching task as found in the final design, with different combinations of 12 words, to minimise the demands of each session and ensure that each word was presented in two out of three sessions. Sessions were completed with a two-day break between. The fourth session then involved the word learning and matching task for all 18 words. In contrast to the final study, perception and production tasks were included as part of the sessions. The fifth session consisted of a battery of tests measuring phonological awareness and working memory, as well as an English vocabulary size test and a background questionnaire. Finally, participants were invited to a short post-test interview over zoom.

Pilot participants were intermediate-advanced learners of English and results from the pilot study showed that participants were able to perceive all the contrasts, although /f-v/ discrimination was lowest and slowest at session four (85% mean accuracy, 593ms mean RT). All participants performed the word learning task with ceiling level accuracy by the third session. Performance in the audio-visual matching task improved over the four sessions and by session four mean performance was above chance for all contrasts. Accuracy was consistently high for /m-n/ words and lower for /f-v/ words, with varying accuracy by spelling presentation. Participants performed at or near ceiling on the production tasks, except for /f-v/ produced with Arabic spelling or no spelling in a picture elicitation task. Interview data was qualitatively analysed and used to develop the language learning strategy inventory in the final design.

Difficulties around recruitment and participation drop out led to the decision that a shorter version of the experiment would be more appropriate, particularly at a time when Covid-19 was a global concern and wanting to minimise the demands on participants' time and energy. It was also decided to move to a post-test questionnaire rather than an interview, as this proved difficult to coordinate in close proximity to study completion, impacting participants' ability to reflect on their experience during the study. However, the interviews proved useful to ensure that participants were able to navigate the design with ease and they reportedly found it a fun experience. In order to reduce the length of the study to a single session of ~30 mins, phonological awareness and working memory tests were removed, as well as the perception and production tasks. Variable quality of audio recordings from production tasks and the inability to control the sound quality for the perception task also motivated this decision.

3.2.2 Pilot study 2

The single session version of the study, with 12 target items differing by two phonological contrasts (/f-v/ and /m-n/), was piloted with 8 L1 English and 6 L1 Arabic participants, and then implemented as the final design. Participants in both language groups were a combination of colleagues, with an advanced understanding of psycholinguistic experiments, and friends without a background in linguistics. This confirmed the function and flow of the experiment, as well as the clarity of instructions and general usability of the platform. Small changes were made to clarify questionnaire wording and ensure easy functionality on a range of different devices and browsers. To increase accessibility, all written questions and instructions throughout the experiment were 'read aloud' by automatic audio recordings. According to pilot participants, this improved engagement and clarity during the study, including those with diagnosed dyslexia. More details about the final design and research instruments are reported in the remainder of this chapter.

3.3 Participants

3.3.1 Recruitment

Following the recommendations of Brysbaert and Stevens (2018), the target sample for both the L1 Arabic target population and L1 English control group was 200 participants each[16].

[16] Ideally 1,600 observations by condition. In the present study there were four trials per 12 items, which varied by three orthography conditions and two phonological contrasts. Thus, eight individual observations by the different conditions, requiring 200 participants.

The L1 Arabic participants could be learners of L2 English with any proficiency level. The L1 English group were required to have little to no L2 Arabic language experience. Recruitment involved three phases. Initially, an opportunity sampling approach was taken, involving recruitment emails being sent out to my personal network in the UK and abroad. The second phase involved contacting language providers and migration support organisations in the local and national area.[17] The third round of recruitment involved posting a video explaining the study in both Arabic and English on my social media channels (Facebook, Instagram, Twitter and WhatsApp)[18]. A snowball sampling approach was also embedded in each phase, as the recruitment information and sign-up link encouraged people to share the study with their personal networks. No financial incentives were offered for participation, as this was deemed inappropriate for participants who may be vulnerable and facing financial insecurity, such as those in complex migration contexts. Additionally, the international scope of recruitment was a complicating factor. Instead, pronunciation videos were created and shared with participants, with a focus on L1 Arabic pronunciation difficulties with L2 English. Additionally, participants were given the option for their score on the matching task to be entered on a leaderboard for additional motivation[19].

Before gaining access to the online study, participants were required to provide their email address on an expression of interest Google form, after which they were sent a generic link and their participant ID to login to the study. The personalised login enabled them to easily re-enter the experiment at the point they got to, if they needed a break or were interrupted for any reason. In addition to the login information, the on-boarding email contained brief information about the aim and procedure of the study, as well as a full copy of the information sheet and consent forms (appendix I and II). Upon completion of the study, an email was sent out with debriefing information (appendix III) and a link to the series of pronunciation tutorial videos I created[20]. Finally, at the end of the project, an accessible summary of findings was provided to all participants (appendix IV).

[17] It was during the second phase that conversations were had between myself and language providers about ethical and practical concerns around recruitment of more vulnerable language learners (Shepperd, 2022).
[18] Twitter recruitment videos: Arabic version: https://tinyurl.com/25kaz3vb and English version: https://tinyurl.com/ycvxebau
[19] Only the top twenty scores were shown, so as not to discourage low-scoring participants.
[20] YouTube pronunciation videos: https://tinyurl.com/56x5dywf

Data collection for the Arabic version of the study was open for three months and the English version was kept open one month longer. In this time, 381 L1 Arabic participants expressed interest in participating, whereas 182 L1 English participants signed up. Of those that signed up, 180 L1 Arabic participants and 145 L1 English participants went on to begin the study. Those who dropped out during the experimental session were not included in the study (L1 Arabic $n = 50$, L1 English $n = 12$). Seven L1 Arabic participants were then removed from further analysis as they did not consent to one or more of the items listed on the consent form. Issues around recruitment, participation rates and informed consent in relation to internet-based research and the present study is further discussed in Shepperd (2022). While the target sample size was not achieved, the number of participants was deemed satisfactory, and indeed a notable improvement, in the context of related studies and data collection during the Covid-19 pandemic.

3.3.2 Background information and exclusions

A total of 128 L1 Arabic participants and 133 L1 English participants completed the session; however, several exclusions were made to address issues of validity and data quality. An overview of the demographic, language background and environmental variables (equipment, distractions, location), before exclusions were made, is available in appendix V, as well as the procedure for data exclusions. Data was excluded for the following reasons:

- **Age**: The study focuses on adult language acquisition and should not include participants under the age of 16.
- **Vision/hearing**: Participants must report normal or corrected-to-normal vision and hearing.[21]
- **Arabic language ability**: L1 English participants were required to have minimal experience with Arabic, particularly literacy in the Arabic script. Therefore, those reporting reading ability above zero in Arabic were excluded.
- **Distractions**: Participants who reported disproportionate levels of distraction were excluded. This was calculated as distraction greater than the mean + standard deviation * 2.5.

[21] As the rate of hearing difficulty in older individuals (aka Presbycusis) increases in adults over 65 years old (Gates & Mills, 2005), participants over this age were excluded from the study.

- **Same day completion**: It was important that there was no sleep cycle between teaching and testing[22]. Therefore, participants that completed the learning phase on a different day to the testing phase were excluded.

Upon reflection, it was decided *not* to exclude on the basis of:

- **Dyslexia/cognitive impairments**: Participants were not excluded on this basis for two reasons, (1) L1 Arabic participants were confused by this question, and (2) dyslexia and other cognitive impairments are not widely assessed, as well as there being more stigma in many Arabic-speaking countries (Aboudan et al., 2011; Al-Qadri et al., 2021), meaning self-report is unlikely to be a reliable indicator.
- **Completion times**: Participants were not excluded based on extreme completion times, as long as they had completed the study on the same day. This took into account interruptions and pressures of participating outside of a lab, where participants were told they could step away from the experiment and return later.

Following this rationale, the final dataset included 114 L1 Arabic participants and 118 L1 English participants, where exclusions resulted in an 11% data loss of both language groups. An overview of the demographic, language background and environmental variables for participants included in the analysis is provided in table 3.1.

Table 3.1: Overview of participant background information

	L1 Arabic			L1 English		
	n	**Mean**	**SD**	*n*	**Mean**	**SD**
Age[1]	114	31.19	7.697	118	35.21	12.353
Gender						
… female	86	75.4%		76	64.4%	
… male	28	24.6%		38	32.2%	
… not listed	0	0%		4	3.4%	
Education[2]						
… primary	0	0%		0	0%	
… secondary	12	10.5%		12	10.2%	

[22] Lexicalisation of newly learned similar-sounding lexical items can occur after only one sleep cycle, where spoken words are integrated into the mental lexicon sufficiently to engage in lexical competition (Dumay & Gaskell, 2012, 2007).

… professional qualification	2	1.8%		15	12.7%	
… bachelors	36	31.6%		38	32.2%	
… masters	43	37.7%		41	34.7%	
… doctorate	21	18.4%		12	10.2%	
L1 dialect[3]						
… American				18	15.3%	
… Australian				1	0.8%	
… British				90	76.3%	
… Canadian				2	1.7%	
… Irish				1	0.8%	
… unknown				6	5.1%	
… Egypto.Sudanic	11	9.6%				
… Mesopotamian	7	6.1%				
… Maghrebi	17	14.9%				
… Levantine	10	8.8%				
… Gulf	69	60.5%				
Self-reported level[4]						
…				1	0.8%	
… none	1	0.9%		113	95.8%	
… beginner	12	10.5%		3	2.5%	
… intermediate	36	31.6%		1	0.8%	
… advanced	52	45.6%				
… nearnative	13	11.4%				
Proficiency test score[5]	114	9.649	2.149			
Audio setup						
… earphones	38	33.3%		32	27.1%	
… headphones	12	10.5%		25	21.2%	
… device speakers	64	56.1%		61	51.7%	
Device type						
… computer	49	43%		82	69.5%	
… mobile	61	53.5%		35	29.7%	
… tablet	4	3.5%		1	0.8%	
% distraction	114	23.5%	27.5%	118	16.5%	23.1%

[1] age is in years
[2] education refers to completed qualifications
[3] regional dialects have a different status for L1 English and Arabic groups, where the dialect for the Arabic group refers to the colloquial variety of Arabic, used alongside MSA.
[4] level of ability in L2 English or Arabic
[5] scale of 0-12

Across both language groups, the sample was skewed towards female, highly educated participants, where a large proportion of participants held a postgraduate qualification. It is likely that the opportunity and snowball sampling approaches attracted participants who already had an interest in linguistics research, which may explain the overrepresentation of postgraduate educated participants. Additionally, the lack of financial incentive may have contributed to attracting participants with such intrinsic motivation to participate, as opposed to a more representative sample. However, this sample does still move beyond the age range and educational experience of typical undergraduate student populations. Additionally, as gender was not expected to be a contributing variable, the predominantly female sample was not a concern.

Dialect corresponded almost exactly to the reported nationality of participants, who completed the study from all over the world. The majority of L1 Arabic participants were from Saudi Arabia (55%), followed by Algeria (12%), Egypt (8%) and Iraq (8%). Arabic-speaker dialects were grouped into the five regional categories outlined in chapter 2.3.1, reflected in table 3.1. Most L1 English participants were from the UK and spoke British varieties of English (76%). A range of proficiency-related variables were measured for level of L2 English and L2 Arabic, for each of the groups. The table reports two key variables of self-reported level and score on a short English test for the L1 Arabic group. The L1 Arabic sample was skewed towards higher levels of English proficiency, reflected in both self-reported data and test scores, but a range of levels were represented. Based on the included measures, it is safe to assume that most, if not all, L1 Arabic participants were also literate in the Roman alphabet.

In order to understand the international contexts of the participants, it was of interest whether individuals participated from their home countries, and if not, where in the world they participated from. Due to the sensitive nature of questions around immigration status, participants were not asked to disclose any such information. This was especially important for the inclusion of people seeking asylum and refugees, who would rightly guard this information carefully or potentially not be sure of their own status. Instead, for speculative insight, participants were grouped by whether their stated location differed from their country of nationality. Most individuals participated from their country of nationality (L1 Arabic = 67.2%, L1 English = 78.2%). However, Syrian participants were all located outside of

Syria, and similarly Palestinian, Sudanese, and Libyan participants were also located outside their countries of nationality, all of which are regions of prolonged conflict and political instability. As mentioned, reasons for being in a different country are speculative, highly complex, and tangential to the current study. The relevance of exploring this dimension relates to the ethical considerations around inclusivity, trauma sensitivity, as well as exploring the language experience of individuals.

Regarding the environmental factors which may have influenced participation in the study, it is noteworthy that over 50% of the participants in both language groups listened to the audio through their device speakers, rather than using headphones or earphones. Furthermore, most of the L1 Arabic participants completed the study using mobile phones, whereas most of the L1 English participants used a computer. A final consideration is that of distractions. From the L1 Arabic participants, one person reported to be 100% distracted, under the influence of alcohol, and provided generally irrelevant responses. Meanwhile, from the L1 English participants, some gave an indication that they were multitasking while completing the study, such as watching TV or having dinner at the same time. This demonstrates the usefulness of collecting this type of data for internet-based research and helps direct choices around exclusion. In the final dataset, the level of reported distraction during the study was lower for the L1 English group (*M*=16.5%, *SD*=23.1%) compared to the L1 Arabic group (*M*=23.5%, *SD*=27.5%). The most common distractions were background noise, interruption from a household member, phone notification or tiredness.

3.4 Stimuli design

As outlined in chapter 2.4, the present study focuses on the acquisition of two different phonological contrasts, /f-v/ and /m-n/. These contrasts were chosen for their phonological and orthographic mappings between English and Arabic. The target contrast /f-v/ is not commonly established in varieties of Arabic, and the assimilation of /v/ to the L1 /f/ category is reflected in the shared orthographic symbol <ف> for both sounds, particularly noticeable in loanwords, such as "video" - <فيديو> - /fiːdjo/. Meanwhile, /m-n/ functions as a control contrast, as it is a well-established contrast in both English and Arabic languages, as well as being consistently represented by two separate letters in each orthography. On this basis, it was predicted that it would be more difficult to lexically encode the distinction between words differing by /f-v/ compared to /m-n/, where the Arabic written forms would accentuate the perceived homophony of the /f-v/ words and the English written forms would visually distinguish the confusable sounds.

The stimuli consisted of 12 bisyllabic CVCVC pseudo-English words, which were matched with images from the NOUN database (Horst & Hout, 2016). Half of the words differed by /f-v/ and the other half differed by /m-n/, where one pair of each phonological contrast was presented in one of three different orthographic input (OI) conditions. The orthographic conditions included the presentation of the auditory form and image with (1) English OI, (2) Arabic OI, or (3) no OI. An overview of the stimuli is provided in figure 3.1. Pseudowords were chosen over real words to manage L1 association, L2 vocabulary differences and potential for ambient exposure outside of the experiment. The number of items were chosen in line with previous single session experiments using pseudowords (Cerni et al., 2019; Escudero et al., 2008; Hayes-Harb et al., 2010, 2018; Mathieu, 2016; Showalter & Hayes-Harb, 2013)

Figure 3.1: Overview of the stimuli, including the /auditory form/, corresponding image and accompanying orthographic input

The first syllable consisted of the target contrast in the context of an unrounded open front vowel /a/ for perceptual salience, as there is evidence that perception of English consonants by Arabic speakers is easier in this vocalic context (Shafiro et al., 2012). The second syllable consisted of two consonants that were found in both English and Arabic, connected by the highly frequent English mid central vowel, /ə/ (a.k.a. schwa). Schwa is the most common vowel sound in English and bisyllablic nouns typically carry the main stress on the first syllable (Guion et al., 2003), thus indicating a fitting option to imitate English word patterns. The CVCVC pattern was chosen as it is common in both English and Arabic (Froud & Khamis-Dakwar, 2021), thus any additional difficulty posed by learning the words is avoided as much as possible e.g., no consonant clusters or difficult vowel sounds.

The pseudowords were generated by WordGen software (Duyck et al. 2004), with additional lists created with Wuggy software (Keuleers and Brysbaert, 2010). Words were generated based on English, following the patten 'target contrast' + a + * + e/i + * (e.g. va*e* and va*i*), excluding the letters <c g h j p q r w x y> due to possible confusion that would distract from the experimental objectives. For example, the letter <c> can be pronounced as [k] and [s] as a single letter or [tʃ] and [ʃ] when combined with <h>, thus highly opaque and would introduce undesirable interference. For another example, the letter <p> was avoided, not due to orthographic opacity but rather because /p/ is not found in the Arabic phonemic inventory, which would again be counterproductive for this experiment. After generating as many items as possible meeting these criteria, minimal pairs were matched together for all the listed items, resulting in 8 minimal pairs for /f-v/ and 12 minimal pairs for /m-n/.

The lists of minimal pairs were recorded by a female phonetician and native speaker of British English. Each word was recorded within the carrier sentences "Here is a…", "where is the…" and as an individual word, relating to different tasks in word learning and testing. Each sentence was repeated three times, one instance of which was then chosen based on clarity and comparability of intonation across stimuli. Recordings were made with a Marantz PMD661MKII handheld recorder and were analysed using PRAAT. Written and auditory forms were normed with three L1 Arabic and three L1 English-speakers, in order to choose the final 12 items. Norming participants completed a listen and spell task and a read aloud task, with the addition of L1 association assessment of both auditory and written forms, all of which is detailed further in the next section.

3.4.1 Auditory and written stimuli norming

Firstly, norming participants listened to a recording of each of the generated words, in a random order, repeated three times by an L1 English speaker, and were asked to write down the best spelling for the word in their L1 script. They were then asked to highlight any words that triggered a strong association with a real word in their L1 from the list. After this, the proposed spellings of the words, in either English or Arabic scripts, were presented in a list and participants were recorded as they read each word aloud. They then completed a questionnaire with the same proposed spellings and were asked whether the pseudoword spelling was close to a real word in their L1. The results from these tasks were explored to see which words had the most consistent sound-spelling correspondences in both English and Arabic and to avoid strong associations with any real words in either language. Any words that were identified as a real word, based on the auditory or written form, were excluded, followed by words with low agreement or accuracy on the listen and spell task and then words with low agreement on the read aloud task.

As mentioned in 2.3.1, Arabic words are formed from roots and patterns, and this may influence how participants engage with the generated stimuli for this study. The pattern of the stimuli in this study, with the main stress on the first syllable containing a open front vowel in a bisyllabic noun, is likely to be perceptually aligned with the pattern of an agent noun, otherwise known as an 'اسم فاعل' ('*ism fa'il*' or active participle). This was indeed the case when norming stimuli with L1 Arabic-speakers, who all consistently transcribed the spoken English pseudowords following the 'فاعل' pattern, or C + /a:/ + C + /i/ + C. When asked, all confirmed that they were certain the first syllable contained a long vowel and recognised the pattern. However, they did not consciously consider the English word in relation to the meaning implied by the agent noun form in Arabic. So, while this pattern in Arabic may bias participants' perception of the vowels, this is unlikely to introduce problems for the present study. Firstly, it is an advantage that the formation of these pseudowords is possible for nouns in both English and Arabic. Secondly, the length of the first vowel and the interpretation of schwa as a short close front unrounded vowel should not compromise the testing of the target consonants. Finally, the perceived familiarity of this word formation in Arabic avoids undesired difficulty during learning.

In Arabic, there was very high consistency between all participants, both when converting the words to text and reading spelled forms aloud. All participants wrote and read words with the structure mentioned above for the agent noun, which was then adopted for all the words presented with the Arabic script (e.g. <ماست> - /ma:sət/- CV:CVC). Predictably, the L1 English-speakers were far less consistent, so commonality was sought between spellings, with adjustments appropriate for the needs of the experiment. Overall, many words were spelled with double letters word-medially during norming, which was not incorporated into any of the final English spellings. It was also found that, for most words, schwa in the second syllable was spelled as <i> and, on the odd occasion, as <e>. A reason for avoiding <e> in the second syllable is the way it can influence the reading of the vowel in the first syllable. For instance, during the listen and spell task /faməs/ was written as <famis>, whereas the written word <fames> was read aloud as /feɪmz/. So, words presented with English spelling followed the pattern: C+ <a> + C + <i> + C.

3.4.2 Images and norming

Related studies using pseudowords have often assigned meanings with black and white line drawings of easily-recognisable objects (Mathieu, 2016; Rafat & Stevenson, 2018; Showalter, 2018; Showalter & Hayes-Harb, 2013) or non-objects (Escudero, 2015; Escudero et al., 2008). In contrast, the present study associated target items with colour photos of novel objects. This decision was based on several factors that arose from designing an experiment accessible to adults with varied educational and cultural backgrounds. Firstly, neurological studies investigating the processing of different types of images by adults with limited literacy have demonstrated that literacy affects visual as well as linguistic processing. There is substantial evidence, for example, that 3D objects are named much faster than 2D objects, such as line drawings (Brucki & Rocha, 2004; Manly et al., 1999; Mansur et al., 2006; Mathuranath et al., 2003; Reis et al., 1994, 2001; Rosselli et al., 1990; Van Der Elst et al., 2006). Additionally, colour and degree of realism increase speed of naming (Reis et al., 2006), leading Bigelow and colleagues (2010) to summarise that "the more lifelike the images, the easier they are to understand or interpret for adults without literacy and formal schooling" (p. 8). Therefore, high resolution colour images that resemble "lifelike" objects were preferable over line drawings on the basis of inclusivity.

Secondly, when deliberating ethical issues around deception and potential for harm for language learners, it was decided that the novel words should be transparently presented with novel objects. While information about the experiment and the artificial nature of the linguistic stimuli was repeated through written and oral instructions in the L1 of participants, this choice restricts the potential risk that learning these novel words could impact real-word language outside of the experiment. For example, learning an invented term for a specific type of dog toy or part of a cleaning device is less disruptive than learning a false label for “a pencil” or another everyday object. Therefore, the constant and explicit ‘novelty’ of both word forms and the meanings they map to avoids misleading participants.

The images for the present study were taken from the Novel Object and Unusual Name (NOUN) Database (Horst & Hout, 2016), which offers high-resolution photos of real objects that vary by colour, texture, and complexity. These images are simultaneously life-like and sufficiently obscure to appear highly novel. Additionally, there is an element of playfulness connected to these images that is exploited to frame the experiment as a learning game, in order to maintain the interest of participants despite the artificiality of the tasks. The database is made up of 64 items, which have been tested for familiarity, name-ability and novelty. Familiarity refers to the percentage of adults who had reportedly seen the object previously. Meanwhile, name-ability refers to the percentage of participants who spontaneously gave the same name to an object. Novelty then refers to the consensus on what objects were and what to call them.

The final 12 images were chosen for this study based on familiarity, name-ability and novelty measurements, combined with norming data from four L1 Arabic-speakers and three L1 English-speakers. Participants named the objects, as well as indicating on a 5-point Likert-scale how they felt about that object (i.e., extremely negative, somewhat negative, neither positive or negative, somewhat positive, extremely positive). This additional norming was conducted to make sure that (1) the assumptions of novelty could be extended to Arabic-speakers who have spent most of their lives outside of the UK, and (2) investigate whether the objects differed in positive or negative association in a way that should be taken into consideration. None of the images elicited any naming agreement between the L1 Arabic or L1 English participants and all the images elicited the average affective response of “neither positive or negative”.

3.5 Procedure

The study was designed as a single session experiment, with the addition of an optional delayed post-test after 24 hours to facilitate insight into developmental processes after a sleep cycle. Due to low take-up and inconsistent timings of the delay, the present study is based on the first session only. The principal elements of the session were a background questionnaire, word learning phase, audio-visual matching task, and a debrief/posttest questionnaire. The experimental flow of the session is provided in figure 3.2. As the figure demonstrates, participants progressed from the information sheet and consent form to an audio check. Subsequently, the L1 Arabic participants completed a short English proficiency test, before moving on to the main experiment. At the end of the audio-visual matching task, participants responded to a distraction check, before completing the post-test questionnaire. The procedure was identical for both L1 Arabic and L1 English participants, except for the English proficiency test. The session lasted approximately 30-35 minutes for the L1 Arabic group and slightly less for the L1 English group, as they did not have to do the English proficiency check[23]. More details on the research instruments are provided in the following section.

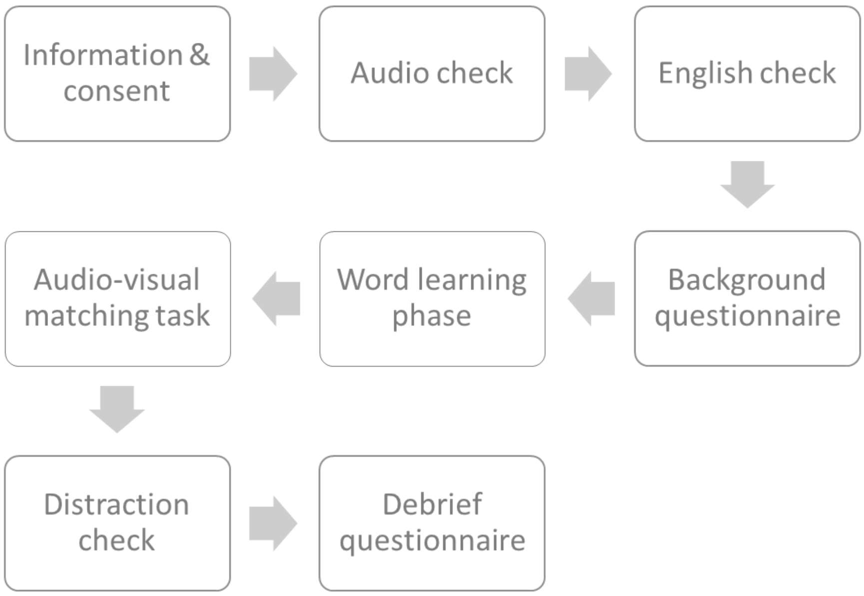


Figure 3.2: Experimental flow of the study session

[23] Timing varied widely, especially as participants could step away and re-join with their login details, so only a rough estimate is given here. The main control for data quality here was the insistence that both word learning and testing took place on the same day.

3.6 Research Instruments

Each component of the experiment is detailed below, with reference to the previous research that inspired the design and implementation in Gorilla experiment builder (Anwyl-Irvine et al., 2020). As mentioned, all written content presented on the screen was 'read aloud' to the participants, both for clarity and accessibility. Arabic audio recordings were made by a female L1 Arabic-speaker from Saudi Arabia, speaking in MSA, who was also a doctoral candidate at the University of York. Meanwhile, I made the audio recordings in English. Recordings were made using either a Marantz PMD661MKII handheld recorder or the PVD Field Recorder 9.7.1 Android app, and were analysed using PRAAT. The only part of the experiment that was not accompanied by this 'read aloud' feature was the information sheet and consent form, as this was not possible to implement.

3.6.1 Audio check

After an initial welcome screen, participants were asked about their audio equipment and it was recommended that they used headphones or earphones, if possible. Upon selecting the relevant image, they progressed to another screen which gave them the opportunity to check their volume. It was also recommended that they stay in a quiet space with no distractions. The audio consisted of a man's voice saying 'hello', set to an intensity of 70dB. Figure 3.3 presents the screens from the audio check in both Arabic and English.

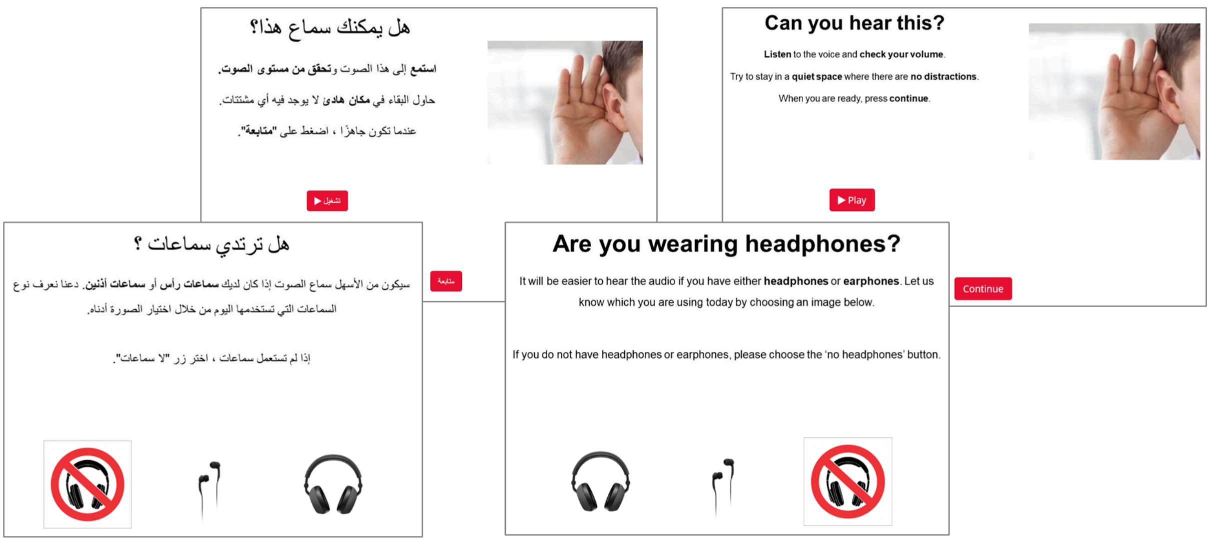


Figure 3.3: Audio check screens for L1 Arabic and English participants

3.6.2 English check

To provide a more objective measure of English proficiency, accompanying the self-reported measures, L1 Arabic participants completed an English proficiency check. This adapted items from the Oxford Placement Test (Allan, 2004), to create a short assessment of grammar, vocabulary, listening, reading, and writing. The first eight questions were gap-fill sentences with multiple-choice answers, as well as the option "لا أعرف" (*I don't know*). Four questions tested grammar and four tested vocabulary, in an order of progressive difficulty. This was followed by a screen with a short recording of the question "what day is it today?", where participants were instructed to listen and type their answer in the text entry box. Finally, participants were asked to read and answer the question "what are your favourite foods?". Example screens are presented in figure 3.4. The questions were intentionally basic and short, to avoid early dropouts or discouragement for lower proficiency participants. While this was a rudimentary assessment of proficiency, it clarified that most participants were literate in English and the high correlation with other self-reported measures demonstrates that this offered insight into English ability (see appendix V). Participant responses were scored out of 12. The gap-fill answers received one point per correct answer, while the two open response questions were given two points, depending on comprehension and the production of a well-formed answer.

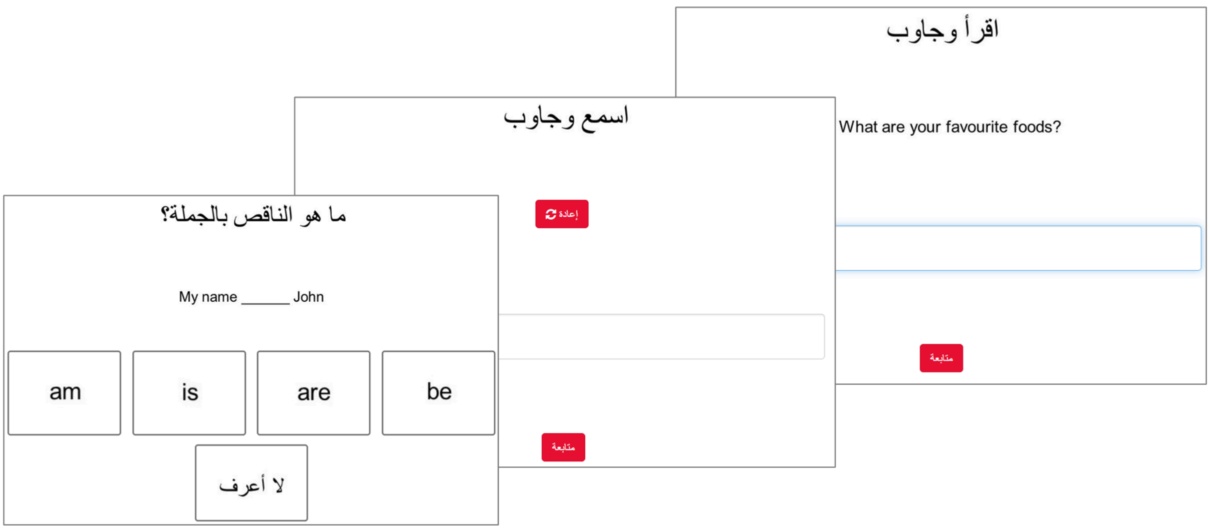


Figure 3.4: Example screens from the English proficiency test, including the grammar and vocabulary gap fill, listen and answer, and read and answer screens

3.6.3 Background questionnaire

The background questionnaire consisted of demographic questions, Arabic language questions, English language questions and exposure to other languages. The background questions included open questions about age and nationality, as well as multiple-choice questions about gender and level of completed education. In line with Bassetti et al. (2020), the following language background questions were added: age of onset of English acquisition; length of English study; length of residence in an English-speaking country; and self-reported amount of English reading, listening, and interaction every day. Additionally, they rated their proficiency in reading, writing, listening and speaking in English, and reported any English language qualifications. Participants also described their L1 exposure and usage. As Arabic is a diglossic language, participants were asked about the colloquial dialect they speak as well as the amount of time reading, listening, and interacting in both dialect and MSA. Finally, they were asked if they used any other languages and how much time they spent reading, listening, and interacting in those languages every day. The questionnaire was the same for both language groups, except L1 English participants were not asked about their English proficiency or qualifications. Additionally, L1 English participants were not asked to distinguish between MSA and Arabic dialects, or asked for qualifications in Arabic. The list of questions featured in the demographic questionnaire can be found in appendix VI. Each question was presented on a separate screen and an example of some screens from the English version are presented in figure 3.5.

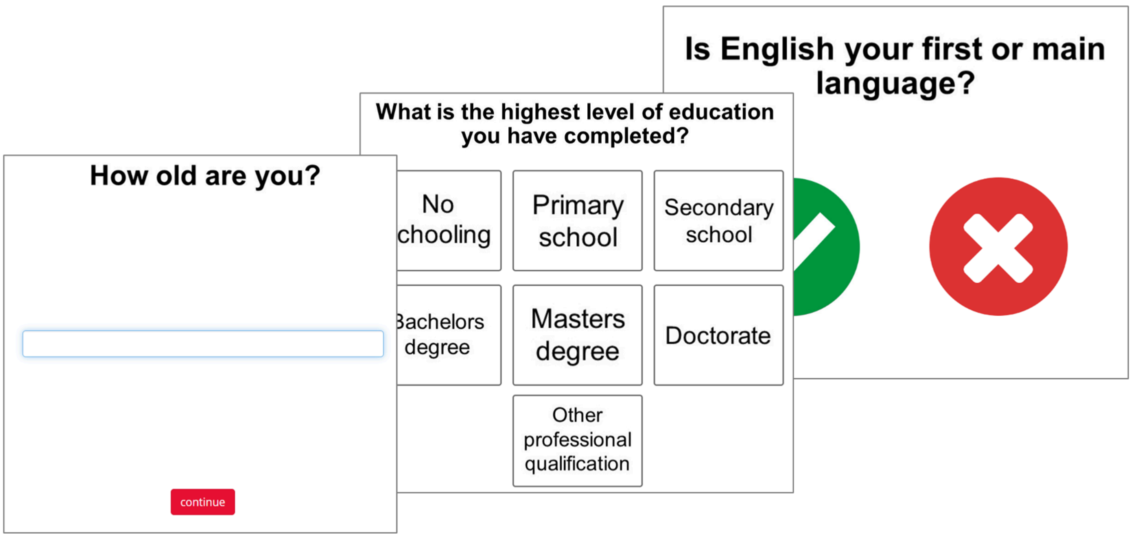


Figure 3.5: Example screens from the English version of the background questionnaire

3.6.4 Word learning phase

The design of the word learning task was based on studies investigating the influence of orthographic input (OI) on L2 learners acquisition of novel words, detailed in 2.2.3 (Bassetti et al., 2015; Escudero et al., 2008; Hayes-Harb et al., 2010, 2018; Mathieu, 2016; Showalter & Hayes-Harb, 2013). In contrast to previous design, the present study integrated testing with feedback throughout the learning phrase, rather than a criterion test after all the words were presented. This decision was made to increase engagement, minimise loss of attention, and reduce dropout rates. The present study adapted the procedure from Rafat and Stevenson (2018), who presented words in blocks of triplets, which were then tested immediately after presentation. Due to the number of stimuli items and combination options, target items were presented in pairs that were non-minimal pair items from the same OI condition. Each word was presented four times over 24 randomised blocks, meaning it was presented twice with each possible paired word. In each block, the pair was presented in a random order, where the participant heard the spoken word, saw the accompanying image, and saw written form (if available) for four seconds (see figure 3.6). The target items were presented in the carrier sentence "here is a _".

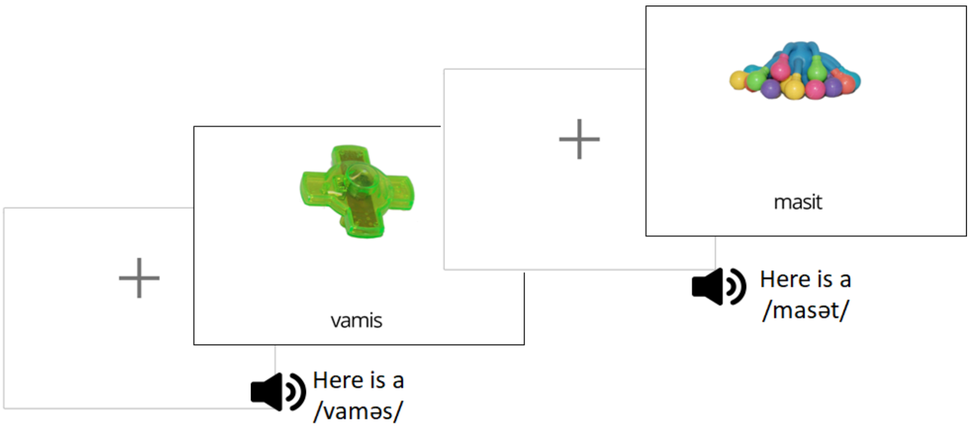


Figure 3.6: Word learning task presentation trial sequence

Immediately after the presentation of each pair, participants moved to testing trials with the pair of words they had just been exposed to. Testing trials consisted of the audio for one of the words they just learned alongside four images on the screen in a grid to choose from (see figure 3.7). Two of the images were the word pair they were just taught and the other two were distractors. The distractors were never a minimal pair item and never started with the same phoneme as the target item. Therefore, the distractors were also never from the same OI condition. They were also counterbalanced so there was always one distractor from each OI condition, meaning that target pairings were not reinforced incidentally by distractors appearing too frequently together. The target words were presented in the carrier sentence “where is the_?”. The second testing trial then played the audio of the other word from the pair and the same images in the grid. The order of the images on the screen was shuffled for each trial and the trial order was randomised. After each response, the participant was given immediate feedback with the appearance of a green tick or red cross accompanied by a sound effect. They were also told their running score after each block of two learning and two testing trials. Before starting the word learning phase, there were two practice blocks, which involved audio and images for the real words “fan”, ”van”, “map” and ”nap”. The use of real words reduced the amount of unfamiliarity participants had to navigate when understanding the requirements of the task.

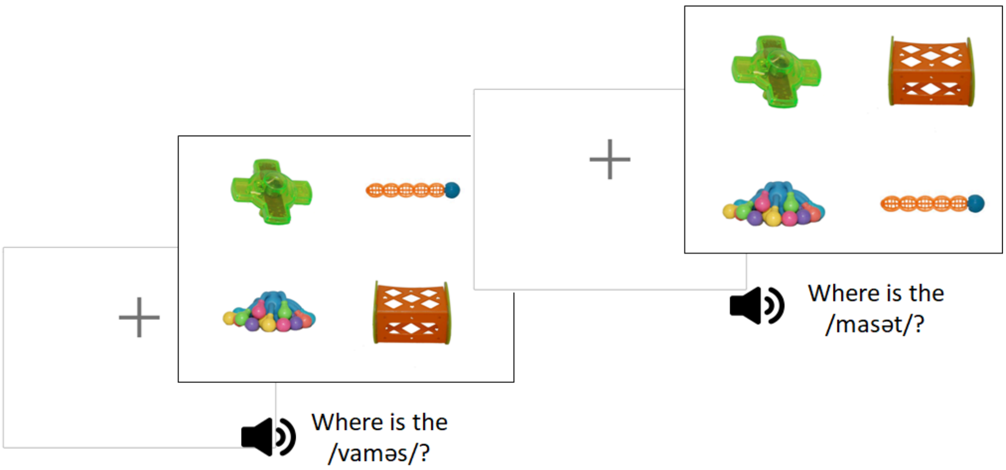


Figure 3.7: Word learning task testing trial sequence

3.6.5 Audio-visual matching task

The audio-visual matching task was a continuation of the designs implemented in the aforementioned studies (Bassetti et al., 2015; Escudero et al., 2008; Hayes-Harb et al., 2010, 2018; Mathieu, 2016; Showalter & Hayes-Harb, 2013). Each word was presented auditorily with either the matching image or with the image of its minimal pair item. Participants then indicated whether the image and the auditory word match, yes or no (see figure 3.8). Each item was presented twice in a match trial and twice in a mismatch trial, resulting in 48 trials per participant. Each trial was preceded by a 500 ms fixation screen and then a screen presenting an image with a green tick and a red cross alongside the target word audio, which was not presented in a carrier sentence. Participants received immediate feedback upon selecting either the tick or cross, through an associated correct or incorrect sound effect. The trials were randomised and split into eight blocks of six trials. Between each block there was a break screen, which also presented participants with their running score. Before starting the test, there were six practice trials with the same real words and images as used in word learning practice trials.

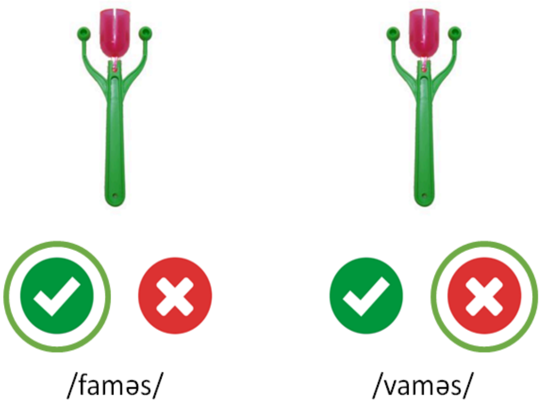


Figure 3.8: Example of a match and mismatch trial in the audio-visual matching task

3.6.6 Distraction check

After completing the learning and testing phases, participants were asked how distracted they were on a sliding scale, from “no distractions” to “a lot of distractions”. The subsequent screen asked participants about the type of distraction, where they were able to choose as many of the 10 options as was relevant, as well as the option to enter their own answer. The options given included: technical problem, phone call, phone notification, alarm, member of household, doorbell, background noise, tiredness, health issue, or not applicable. Figure 3.9 presents the screens from the distraction check in both Arabic and English.

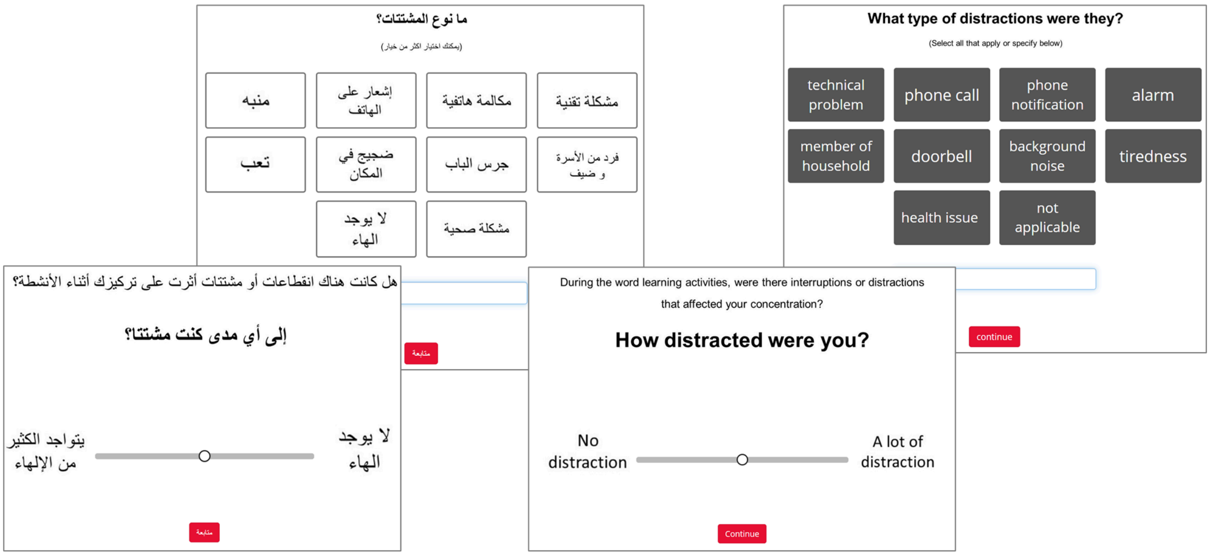


Figure 3.9: Distraction check screens for L1 Arabic and English participants

3.6.7 Post-test questionnaire

The study concluded with a reflective post-test questionnaire, consisting of three main parts: awareness of the influence of written input; speculative spelling; and a language learning strategies (LLS) inventory. All the questions, in English and Arabic, can be found in appendix VII. The questionnaire items were formed from pilot interviews and questionnaire responses from a similarly designed study for my Master's dissertation. The first part of the questionnaire contained eight open and multiple-choice questions to investigate participants' awareness of exposure to the different types of written input, and their perception about how that influenced their learning of new words. Then, questions related to seeing the words with Arabic, English, or no written input investigated predictions as to whether one OI condition facilitated learning over the others. Next, participants were asked about learning new words with written input in general, to gain insight into broader beliefs and practices around language learning and written input. The second part asked participants to write a speculative spelling for the four words that were presented with audio-only. This was another way to assess how well participants encoded the contrast between minimal pairs and it was open for them to respond in English or Arabic letters.

The third part consisted of an open question about how participants tried to learn the words during the study, followed by 23 statements about LLS. One of the most widely used and tested questionnaires designed to assess LLS is the Strategy Inventory for Language

Learning (SILL) (Oxford, 1990). LLS research has expanded greatly over the past three decades, including adaptations of SILL to focus more specifically on areas such as pronunciation. The statements included in the present study have evolved from adaptations of SILL statements and modified versions of SILL focused on pronunciation strategies (Berkil, 2008; Peterson, 1997; Rokoszewska, 2012). Due to the limited research into phonology and orthography, the strategies reported in my master's dissertation and piloting of the present study, alongside pronunciation strategy inventories, were used to devise what is henceforth referred to as the *Phonology and Orthography Language Learning Strategy* (POLLS) inventory. The reason for developing new strategy statements was further motivated by the actual-task strategy questionnaire design, meaning learners were asked about strategies used in a recently completed task, as opposed to broader reflections on learning. As such, many of the statements from broader strategy inventories about vocabulary learning, pronunciation and reading were not relevant to word learning in an experimental context. The next section outlines the development of the POLLS inventory.

3.6.8 Development of the POLLS inventory

Questionnaire responses from 35 L1 English adults, who participated in a similarly designed study for my master's dissertation project, were revisited in order to develop relevant strategy statements for the present study. Responses to an open question, asking about how participants tried to learn the words during the study, were qualitatively analysed through a process of open and axial coding. The direction of analysis was to discover categories related to the strategies participants used to learn new words with different types of written input in a word learning experiment. In contrast to the present study, this experiment had a between-subjects design, where participants were divided into three groups which learned pseudo-Arabic words differing by nonnative contrasts with either Arabic spelling, English spelling, or audio-only. Responses broadly fell into cognitive and metacognitive strategies, aligning with Oxford's strategic self-Regulation (S2R) model of language learning (2011, 2016) and other LLS research discussed in 2.1.3.3. Oxford highlights the flexibility of strategy categories and how strategies can be combined into 'strategy chains' that work together. Both observations were evident when analysing the questionnaire responses and should be taken into account as the following categories are discussed in more detail.

Within cognitive strategies, the categories included: *associating*, *analysing*, *contrasting*, *grouping*, *visualising*, *moving*, *mouthing*, and *repeating*. Within metacognitive strategies, the categories included: *recalling*, *focusing*, and *reflecting*. These categories were analysed considering the strategies outlined in Oxford's S2R model (2011, 2016) and an overview of the analysis can be found in table 3.2. The specificity of the strategies mentioned in relation to novel word learning and the influence of written input meant that strategies did not always fall neatly into S2R categories of LLS or fit well with previously constructed statements. This underscores the reason for creating new statements that were not present in any of the inventories mentioned so far.

Table 3.2: Qualitative analysis of master's dissertation questionnaire responses when asked: "What strategies did you use to try to learn the words?"

Type	S2R category	Category	Definition
Cognitive	Conceptualising with detail	Associating	Associating words to known or familiar sounds or words to remember them and/or strengthen connection between sound and imageable meaning.
		Analysing	Breaking words down into smaller units, including focusing on first and last sounds/letters, and looking for patterns across words and sounds in general. (Often used with associating)
		Contrasting	Detecting differences, comparing, and contrasting in order to understand and remember new words, especially similar sounds. (Often used with grouping and associating)
	Conceptualising broadly	Grouping	Remembering words in groups, often based on similar sounds (often used with contrasting and associating)
	Using the senses for understanding and remembering	Visualising	Using mental images, imagining additional visual context, picturing themself with the object. This can involve dynamic aspects such as mentally animating the object and imagining scenarios. (Often used with associating and analysing visual input)

		Moving	Moving or thinking about an action to remember the new words. Other tactile/kinaesthetic strategies apply, in terms of how texture/touch and the body broadly is used to remember and understand words.
		Mouthing	Mouthing, in the sense of moving the mouth to silently form/rehearse the shapes required to produce the sounds of the word. (Often but not always used with repeating)
		Repeating	Repeating out loud or 'in my head', including ideas connected to vocalising connections and using auditory rehearsal of the new words to remember. (Often used with mouthing and associating)
Metacognitive	Monitoring cognition	Recalling	Purposefully trying to recall a word from memory before hearing it again, through mouthing, saying, or thinking of it. This is to create an opportunity for practice, to 'test' themselves, and/or monitor progress. Some reversed this by closing their eyes, to then listen and recall the image and open their eyes to check. (Often used with visualising, mouthing and repeating)
	Paying attention to cognition	Focusing	Directing focus or attention towards or away from elements of the input, including consciously avoiding spelling or sequencing attention to different visual elements.
	Evaluating cognition	Reflecting	Making evaluative judgements about how successful the strategies were. Also includes reflecting on learning and task performance.

To inform the creation of the required statements, further qualitative analysis was conducted with the questionnaire data from my master's dissertation combined with pilot interview data from the present study. The three interviews also mentioned strategies that fell into the categories of *associating*, *repeating*, *analysing*, *contrasting*, *grouping*, *focusing,* and *evaluating*. Next, the findings from the pilot interviews and master's questionnaire were compared to the pronunciation strategy inventories mentioned above (Berkil, 2008; Peterson, 1997; Rokoszewska, 2012). Table 3.3 outlines the categories and example strategies that arose, how they mapped to existing statements in previous strategy inventories, and then the final wording used in the POLLS inventory. Priority was given to keeping the inventory as brief as possible and focused on understanding how learners responded to written input when learning the new words, rather than a comprehensive overview of all strategy usage during the experiment. Further discussion of the research context relevant to the development of this tool and the qualitative approach taken when analysing the data in the present study is provided in chapter 6.

Table 3.3: Comparison of strategy categories and statements for the POLLS inventory

Strategies from qualitative analysis and example	Examples of existing statements	Final statement
Cognitive		
Associating: I thought of sounds and words that could be related to English or French. / I associated them with words I already knew, so the ruler was Kaylum, the name of someone I knew.	I think of relationships between what I already know and things I learn in English (Oxford, 1990). / I create associations between new associations and what I already know. (Peterson, 1997).	I created associations with words or things I already know.
Associating: the word for computer mouse sounded like 'careful', so I remembered careful, don't tread on the mouse.	I memorise the pronunciation of a given word by putting it in a context (a sentence, a story, a rhyme etc.) (Rokoszewska, 2012).	I put the word in a context to remember it (a sentence, a story, a rhyme).
Visualising: One word [for belt] sounded like Gollum to me so I tried to imagine Gollum wearing a belt. / I made an image in my mind of the words.	I remember the word by making a clear mental image of it. (Peterson, 1997). / I memorise the pronunciation of a given word by associating it with an image or picture (Rokoszewska, 2012).	I made a mental image or imagined additional connections to help me remember.
Visualising: I made up a spelling in my head to help me remember the words	I visualise the spelling. (Peterson, 1997). I memorise the pronunciation of a given word by visualising its	I visualised the spelling of the words in Arabic in my mind.

	transcription (Rokoszewska, 2012).	I visualised the spelling of the words in English in my mind
Repeating: I tried to repeat the words in my head to rehearse them. / I said the words out loud to myself through the experiment.	I repeat (out loud or silently) after my teacher… (Peterson, 1997). I repeat aloud after tapes, television...(Berkil, 2008).	I repeated the words out loud or in my head.
Mouthing: I imagined how I would make the sound with my mouth shape. / I tried to mouth the sounds.	I notice a teacher or native speakers' mouth positions and pronunciation and try to imitate them. (Peterson, 1997)	I thought about or practised mouth positions to pronounce the words.
Moving: I wanted to do an action to each word, so for some I could feel like I wanted to move or think about a particular part of my body.	I physically act out the new word. (Oxford, 1990).	I thought about an action or movement to help remember the words.
Analysing: I broke the syllables up into English word sounds./ I tried to work out the separate sounds.	I seek specific details in what I hear or read (Peterson, 1997).	I broke words down into syllables and sounds.
Analysing: I used the written forms to try and find a pattern in letter structure with the sound.	I look for patterns in the new language (Peterson, 1997)). I try to find patterns in English (Oxford, 1990).	I tried to find patterns in the new words and sounds.
Analysing: I tried to remember the first syllable of each word/ I tried to remember that the word for banana ended with an "f".	-	I tried to connect sounds and letters. I tried to remember the first or last sounds.
Analysing: I tried to find any similarities with English and tried to find any patterns with the sounds. / the written forms didn't always	I look for similarities and contrasts between the new language and my own (Peterson, 1997).	I looked for similarities and contrasts between the pronunciation of the

match up with the sounds like English.		new words and the languages I know.
Grouping: After hearing all the words a few times, I began grouping similar sounding words.	I group words that sound similar in order to memorise their pronunciation (Rokoszewska, 2012).	I grouped similar sounding words.
Contrasting: I tried to learn the difference in pronunciation/sounds between groups of similar words.	-	I used the English spelling to distinguish between similar sounds. I used the Arabic spelling to distinguish between similar sounds.
Metacognitive		
Focusing: I tried to keep my eyes off the written words, but this was a discipline. / I couldn't read them, so I focused on the picture rather than the Arabic words.	I decide in advance to pay special attention to specific language aspects. (Peterson, 1997).	I purposefully ignored the English spelling. I purposefully ignored the Arabic spelling.
Recalling: I tried to recall the words before I heard them to practise remembering them. / I tried to test myself by saying them out loud before hearing the spoken voice (to confirm I was correct).	I think about my progress in learning English (Oxford, 1990). /I evaluate the general progress I have made in learning the language (Peterson, 1997).	I found ways to test my memory and recall the new words. I thought about my progress in learning the new words.
Reflecting: I learned a lot about how I learn. My first instinct was to focus on the visual and language structure, but actually for memory retention, that was not helpful at all...I learned I have to hear myself repeating it.	I notice my English mistakes and use that information to help me do better (Oxford, 1990).	I noticed my mistakes and used that information to help me do better.

Affective		
Evaluating: [Arabic OI] makes it more confusing. / I found it easier with the [Roman OI] than with the sounds alone.	I noticed if I am tense or nervous when I'm studying or using English (Oxford, 1990)./ I analyse my feelings connected with learning pronunciation (Rokoszewska, 2012).	I noticed that I felt more relaxed or confident when I saw the Arabic spelling. I noticed that I felt more relaxed or confident when I saw the English spelling.

3.7 Ethical considerations

This study was ethically approved by the Department of Education's ethics review procedure. The potential for harm arising from participating in this study was low, and where possible, steps were taken to improve accessibility, reduce stress and create a study that was enjoyable to participate in. This included the implementation of 'read aloud' text, frequent breaks, ease of access entering and leaving the study, and using a game-like design with feedback throughout. The inclusion of participants with low L1 literacy was included in the original design; however, for numerous ethical and practical reasons, this was not deemed appropriate[24]. As mentioned in section 3.3.2, it was likely that some participants had a background of forced migration and potential for experience with conflict-related trauma. For this reason, additional sensitivity was taken when designing questionnaire items, stimuli and overall procedure, to avoid unintended stress (e.g., avoiding sensitive or triggering material, easy functionality, providing breaks and encouraging messages). The potential for deception related to the learning of an artificial lexicon was mitigated by using novel images to accompany the pseudowords, alongside clear instructions. Additionally, my ability to communicate in written and spoken Arabic facilitated my ability to check the wording of translations and respond quickly to email communications from participants, trouble-shooting technical issues, amongst other things. Attention to smooth functionality across devices was also motivated by accessibility and reducing the possibility of stress during participation.

[24] See Shepperd (2022) for more detailed discussion around barriers to low-literacy participants.

Information, consent and debriefing were all communicated in the first language of participants by email, for a long-term record. The lack of in-person communication meant participants could not ask questions and I could not assess understanding of the information provided. However, internet-based methods offer the benefits of participants being able to read information in their own time and withdraw easily, if they wish to do so. A conservative approach was taken to data protection, where any negative response on the consent form resulted in the deletion of all data from that participant, even if overall consent was given before beginning the experiment. Data reported on in the rest of the study is fully anonymised and non-identifiable, with no sensitive data collected. Anonymised raw data will be made available on OSF, as soon as possible.

No financial incentives were used in this study, as mentioned in section 3.3.1. This proved problematic for including participants from different backgrounds, who may not be motivated by an existing interest in linguistics research. Furthermore, "the idea that people should contribute to research out of personal interest or an altruistic support for the advancement of science is a privileged concept that assumes an affluence of time, stability, and capacity" (Shepperd, 2022, p. 10). As an alternative, participants were given access to tailored pronunciation videos, although views indicated low-uptake. There was also the opportunity to see whether they could score in the top twenty of a public leaderboard, aimed to motivate attention and add to the game-like quality. However, some participants reported being disappointed after not scoring highly, which was an adverse consequence, particularly for participants who already have low confidence learning languages. The inability to communicate with participants directly, such as clarifying misinterpretations of task performance, is a limitation of internet-based research. This underscored the importance of the debriefing information, although it is not possible to know whether participants engaged with these emails. Other considerations related to online research are detailed in the next section.

3.8 Internet-based factors

To conclude this chapter, certain elements pertaining to internet-based research methods deserve closer attention, before moving on to the analysis and results of the present study. In particular, it is pertinent to understand more about the chosen experimental platform, the presentation of stimuli, issues around recruitment and retention, and data quality. Each of these points is discussed below and provides useful context for methodological decisions reported above, as well as subsequent analysis and discussion.

Several online experiment builders were trialled in the transition to an internet-based approach, including OpenSesame (Lange et al., 2015; Mathôt et al., 2012), PsychoPy3 (Peirce, 2007), jsPsych (de Leeuw, 2015) and Gorilla Experiment Builder (Anwyl-Irvine et al., 2020)[25]. The reliability, flexibility and ease of Gorilla made it the optimal choice for the present study. Gorilla is set up to easily build, host and recruit for a wide range of behavioural experimental designs, without the need for coding. Furthermore, Gorilla offers a 'build for free, pay per participant' business model, which keeps costs low. While it is not an example of free, open-source software, Gorilla also promotes open materials on its platform. In addition to the practicalities of building, hosting, and recruiting, it is also important to consider the quality of data when running experiments online. Firstly, reliability of stimuli presentation times and response times is a recurring issue. Secondly, while there is the possibility to increase the scale of a study, online experiments are likely to exhibit higher dropout rates. Thirdly, the loss of control over who participants are and the experimental conditions outside of a laboratory setting need to be considered. I address each of these points in turn and discuss how the present study approached each challenge.

3.8.1 Stimuli presentation and response times

It is intuitively concerning to those accustomed to lab-based experiments that online research introduces numerous opportunities for undesirable variation, which could potentially undermine the validity and reliability of a study. In computer-based laboratory research, timing tends to be relatively reliable and straightforward to report for the following reasons. With regards to presentation, you do not have to contend with varying internet speeds, browser compatibility, operating systems, screen size or screen refresh rates, as your experiment runs through downloaded software on one or multiple equivalent devices. Regarding response times, by using the same hardware and software, there is no need to consider the difference between desktop, laptop or Bluetooth keyboards and mouse responses, as well as touch screen responses on various smart devices. As the present study

[25] For further comparison of tools for online research see Sauter and colleagues (2020).

investigates both audio-visual exposure and response times, these were highly pertinent considerations.

Two recent large-scale studies addressed these concerns by investigating visual and auditory stimulus timing and response times across a range of platforms, operating systems, browsers and hardware (Anwyl-Irvine et al., 2021; Bridges et al., 2020)[26]. Both refer to the distinction between accuracy and precision, where accuracy is the distance from a value of zero error, and precision is related to how variable errors are or noisy the data is. For example, while you ideally aspire to high precision and high accuracy in experiment timings, it is preferable to have high precision and low accuracy compared to low precision and high accuracy, as a consistent delay can still reveal a reliable difference or effect. This means that issues with accuracy, often related to hardware or physical setup, can be measured and corrected, which is not the case for poor precision. An overview of both studies can be found in table 3.4. It goes without saying that the development of online research tools has advanced considerably since the publication of these two studies; however, they provide insight into the state of the technology at the time of design and data collection for the present study.

Table 3.4: Overview of software and hardware tested for timing by large-scale studies

	Bridges et al. (2020)	**Anwyl-Irvine et al.** (2021)
Platforms/Packages	Lab: PsychoPy, Eprime, NBS Presentation, Psychophysics Toolbox, OpenSesame, Expyriment, Online: PsychoPy, Gorilla, jsPsych, Lab.js and Testable	PsychoPy3, Gorilla, jsPsych, and Lab.js
Operating systems	Windows10, macOS, Linux	Windows 10, macOS
Browsers	Firefox, Chrome, Safari, Edge	Firefox, Chrome, Safari, Edge
Hardware	High-performance button box	Desktop vs. integrated laptop keyboard on PCs and Macs

[26] It is noteworthy that Anwyl-Irvine and colleagues are developers with Gorilla, while Bridges and colleagues are the authors of PsychoPy.

Best online options	PsychoPy and Gorilla	No standout best platform but the best OS/browser combination was Windows/Chrome.
Problems to consider	Limitations presenting audio-visual stimuli online. Variability between browser and operating system combinations.	Platform contributes to greater variance than device, but both introduce inconsistencies, as well as operating systems.

These complementary studies describe the complex timing considerations required for conducting precise behavioural research online but taken together offer encouraging evidence that most popular online platforms offer reasonable accuracy and precision. This aligns with an earlier study by Miller et al. (2018) which found internet-based response times were not compromised in comparison to offline laboratory settings. However, both studies emphasise the lack of generalizability of their findings and the need to pay specific attention to the combination of platform, browser, operating system, and devices used within each individual study. The issue of generalisability relates to the point about rapid advances in technology, meaning that the specifics of these studies quickly become outdated. Mobile device usage is increasing, browser functionality is improving and operating systems are constantly being updated, to name a few obvious examples. That is to say nothing of the fact that both studies were conducted before the Covid-19 pandemic, which has led to an explosion of both internet usage and technological development.

The present study addresses these considerations through inclusion of factors related to participant software and hardware within exploratory analysis (appendix V and VIII), and open access data including these variables available for future analysis. Gorilla is pre-programmed to collect trial onset timing, device type, operating system, browser, monitor size and loading delays. It is not possible to know how visual selections are made, as keyboard presses are not used in the present study, but it is assumed that tablet and mobile devices require touchscreen responses, while laptops use an integrated mouse pad and desktops use a wired or wireless mouse.[27] General timing guidance provided through

[27] It is of course possible that laptops also have touchscreen capabilities, a wired or Bluetooth mouse, and/or connection to an external monitor. All of which may influence timings to differing extents.

Gorilla's website (*Gorilla Experiment Builder*, n.d.) indicates that a keyboard or mouse is typically polled every 8ms, while a touch screen on a mobile device is closer to ~20ms. Encouragingly, Pronk et al. (2020) found reasonable timing accuracy for both touchscreen and keyboard devices, indicating there is no evidence that touch screens would prove problematic or vary dramatically from findings reported in other studies with keyboards. Overall, there is good justification for the use of internet-based methods and Gorilla as an appropriate choice of platform in relation to the presentation and response timing requirements of the present study.

3.8.2 Recruitment and retention

Very few studies report or analyse dropout rate, even though some studies report up to 69% dropout rate and there is evidence that dropout rate can interact with experimental conditions (Sauter et al., 2020). Online surveys recruiting through internet-based ads can report dropout rates up to 80%, with an average of around 30%, in contrast to 15% through direct recruitment of individuals and less than 5% for in-person and telephone interviews. Length of survey, level of reported interest and burden are all reported to have a strong effect on the risk of dropout (Galesic, 2006). These findings are not unique to online surveys but extend to most online data collection, including experiments like the present study. Jun et al. (2017) collected data based on 7,868 participants of three short experiments (5-10 mins) and found that 20% of participants who started the experiments did not finish, with the highest dropout rate (38%) found in the longest experiment (10 mins).

Aside from length, it was found that participants motivated to respond out of boredom were the most likely to drop out, whereas those motivated by contributing to science were the least likely to drop out. Similar motivation types are reported in studies recruiting through platforms such as Amazon Mechanical Turk, demonstrating that intrinsic motivation remains an important factor, even when a financial incentive is available (Antin & Shaw, 2012; Kaufmann et al., 2011). Additionally, Sauter et al. (2020) asked 103 Germans on an online opinion poll at what point they would leave an online experiment and found that 44% reportedly would leave after 15 minutes, 35% after 30 minutes, a further 10% after 45 minutes and an additional 12% after an hour. Therefore, it was clear that online recruitment and retention would likely be a sizable challenge for the present study.

Several steps were taken to address these challenges. Firstly, to a) create a more personal tone to aid recruitment, b) promote accessibility, and c) improve informed consent, I recorded a two-minute video outlining the study in both Arabic and English, as mentioned in section 3.3.1. Reminder emails were also used to encourage participation. A further priority was to make the experiment as game-like as possible and compile several short tasks, averaging between 5-10 mins each, to maintain interest and attention. This was important considering the additional task load of a ~30 minute experimental session, as opposed to a 10 minute survey. Tasks were close adaptations of previous lab-based studies, but differed both in length (number of trials) and quantity of feedback. Participants were given feedback for almost all responses, as well as a running score between blocks, and a progress bar for each task. As has been mentioned, incentives were implemented with limited success and would be adjusted for future research. That being said, ~260 people completed the experimental session, and informal feedback indicated that most participants found it a fun and interesting experience.

3.8.3 The 'who', 'where' and 'how' of participation

Research into the reliability and validity of internet-based research compared to laboratory studies have long questioned the limitations of larger, more heterogeneous samples in less controlled conditions. However, even in early online research, the data that arose from internet-based research appeared to be comparable to lab-based studies (Birnbaum, 2000; Mason & Suri, 2012; Peyton et al., 2022; Vaughn et al., 2018). Many have heralded the opportunity to recruit samples that are more representative of the general (online) population, rather than the Western Educated Industrialised Rich and Democratic (WEIRD) samples that dominate lab-based research. Yet, there are several additional considerations that need to be made for successful online data collection with more diverse samples (Rodd, 2019; Sauter et al., 2020; Shepperd, 2022).

Firstly, participants may be less familiar with behavioural experiments, compared to samples drawn from undergraduate students (Andringa & Godfroid, 2020; Plonsky, 2023). Confusion may be exacerbated as the researcher is not available to check understanding or answer questions directly, heightening the importance that instructions are clear, concise, and comprehensible to a wide range of backgrounds. In the present study, all instructions were presented in short sentences, written in the L1 of participants, and accompanied by a recording reading the same sentences aloud. Where possible, images were used to clarify instructions further and all tasks had practice trials before the main task.

Secondly, it can be difficult to ascertain whether participants are who they say they are and who you are looking to recruit (Rodd, 2019). The lack of financial incentives reduced the risk of 'fake participants', where experiments are (1) completed by the same person multiple times, (2) someone who does not fit the inclusion criteria or (3) a bot set up to complete multiple experiments and reap the financial reward. For inclusion in the present study, participants needed to either be L1 Arabic-speakers, who also identified as being learners of English, or L1 English-speakers with Arabic experience. Several checks were embedded within the task design to ensure data quality on this front. For example, alongside self-reported proficiency in both Arabic and English in the demographic questionnaire, participants would not be able to engage with email communication or follow instructions without a high ability in their reported first language. This was further aided by the English proficiency check and the use of open response questions at various points, which participants would not be able to skip through.

Another concern is that participants would be less committed when there is no experimenter observing their progress and that level of attention could affect data quality. However, research comparing attention and desirable responses in domestic and laboratory settings find little evidence to suggest that this is in fact the case (Clifford & Jerit, 2014; Gould et al., 2015; Hauser & Schwarz, 2016). As mentioned, several features were added to make the experiment as engaging as possible, both to improve retention and attention. This also motivated the inclusion of the distraction check. Furthermore, short tasks, regular breaks, and the ability to easily return to the same point later using the personalised login aspired to be flexible around different disruptions. Beyond this, it is very difficult to control the environment in which participants are completing the tasks and how well they are engaging with the experiment (Rodd, 2019).

Finally, as mentioned in 3.8.1, hardware influences how participants experience an experiment and data quality. In addition to timing of presentation and responses, it is necessary to extend considerations to the quality of presentation and responses. In a lab-based experiment, experiments are often completed in a sound attenuated booth with consistent, high-performing hardware and software running the experiment. This hardware also ensures reliable, high-quality presentation of auditory and visual stimuli and auditory recordings. Aside from the deficiencies of JavaScript to deliver precise audiovisual synchrony, participants are likely to have vastly different set ups with regards to devices, screens and speakers/headphones. This motivated the inclusion of the audio check, both for volume control and to find out what equipment participants used.[28] As the present study emphasises inclusion of diverse samples, it was decided to include the self-reported information about audio equipment for later analysis, rather than exclude participants without the ideal equipment. Similarly, participants are not excluded on the basis of device type, even though the difference between screen size of mobile and desktop devices is substantial. In support of this decision, there is evidence that learning and object recognition is largely size invariant (Furmanski & Engel, 2000), such that there is no clear advantage for learning and recognition when an object is larger. As the present study emphasises global recognition of easily distinguishable images, validity is unlikely to be undermined.

In conclusion, a detailed account has been made for the decisions taken during the transition to online data collection. An increasingly strong case is made through rigorous investigation of internet-based experiments across a range of disciplines, providing evidence of data that is of sufficient quality to rival laboratory research. This is further supported by larger and more 'realistic' samples, facilitated through ever-improving technology. The present study has endeavoured to follow recently emerging ideas of 'best practice' at every stage. However, even with the boom in internet-based research over the last decade, the circumstances in relation to COVID-19 were complex, unknown, and incessantly changing. Therefore, in order to respond appropriately, both to the circumstantial and psychological limitations of the researcher and those who participated, the core values of rigour and transparency that guide this work were combined with creativity and flexibility to adapt to these unprecedented times. This should be taken into account in relation to the methodology outlined here, as well as the analysis and results, to which the next chapter turns.

[28] That being said, when listening to auditory elements of the experiment, even if participants all used headphones, audio quality would still have varied between brand, earphones vs. headphones, wired vs. Bluetooth etc.

3.9 Analysis overview

Primary research question: Does exposure to the Arabic and English script differentially influence L1 Arabic-speakers' acquisition of new words that differ by a nonnative phonological contrast?

The primary research question will be discussed further in chapter 8, when drawing together the results from the sub-questions, which are outlined in more detail in the following chapters. Chapter 4 addresses the influence of written input on the lexical encoding of nonnative contrasts, as assessed by the audio-visual matching task. This is followed by quantitative and qualitative analysis of the questionnaire responses to explore learners' perceptions and beliefs around the influence of written input when learning the target items in chapter 5. Next, in chapter 6, a combination of factor analysis and qualitative content analysis are applied to questionnaire responses related to language learning strategies. These findings are then explored in relation to performance on the audio-visual matching task, in order to assess whether different language experience and approaches to learning predict task performance. Finally, in chapter 7, the results are drawn together to explore the role of proficiency in relation to orthographic influence.

Chapter 4: Influence of written input on lexical encoding of contrasts

4.1 Introduction

This chapter presents the analysis of participant accuracy and response time data from an audio-visual matching task, designed to test whether participants' ability to lexically encode the distinction between pseudowords differing by a confusable L2 contrast is influenced by exposure to orthographic input (OI) during word learning. The target phonological contrast is /f-v/, while /m-n/ functions as a control contrast. The data from the two language groups are analysed, firstly L1 English-speakers with little to no experience in Arabic, and secondly L1 Arabic-speaking learners of English with varying L2 proficiency. The English-speaking group functions as a control for reliability and validity, whereas the Arabic-speaking group is the focus of the analysis outlined below. The research question being addressed is:

To what extent do different types of written input influence the lexical encoding of L2 phonological contrasts in memory during novel word learning?

As mentioned in 2.2.3, four key variables reportedly mediate the influence of written input: systematicity, familiarity, congruence and perceptibility (Hayes-Harb & Barrios, 2021). These variables relate to the above research question and present study design in the following ways:

1) **Systematicity**: both /m-n/ and /f-v/ are systematically represented by the L2 English writing system (i.e., <m>, <n>, <f> and <v>).
2) **Familiarity**: L2 graphemes are not familiar to the learners from their L1, but become increasingly familiar with L2 (literacy) proficiency.
3) **Congruence**: For words that are transliterated into the Arabic script, L1 Arabic and L2 English employ comparatively congruent grapheme-phoneme correspondences for /m-n/ but not /f-v/ (i.e., /m/ - <م>, /n/ - <ن>, /f/ and /v/ - <ف>).
4) **Perceptibility**: L1 Arabic-speaking learners of L2 English are unlikely to have difficulty with words differing by /m-n/, as this is an established L1 contrast and

assumed to assimilate to two separate phonemic categories. Meanwhile, words differing by /f-v/ are likely to be difficult, as this is a nonnative contrast, where both sounds are assumed to assimilate to a single phonemic category of /f/.

In summary, words differing by both /m-n/ and /f-v/ in the present study are represented systematically in L2 English spelling, which is deemed unfamiliar on the basis that it employs a different script from L1 Arabic. However, familiarity is logically relative to L2 (literacy) proficiency, thus better articulated as a distinct script from the L1. The primary differences between the contrasts relate to congruence and perceptibility. The control contrast /m-n/ is both congruent in its transliterated form, as well as an easily perceptible native contrast. Meanwhile, the target contrast /f-v/ is neither congruent in its transliterated form, nor easily perceived, as the perceptual assimilation of both /f/ and /v/ to a single category of /f/ is reflected in the transliteration of both sounds using one letter <ف>. Therefore, it is predicted that the systematic representation of the target words in L2 English script input will have a different influence on Arabic-speakers' ability to lexically encode the nonnative /f-v/ contrast compared to the incongruent representation of the target words in the transliterated Arabic spelling. The following section briefly returns to the findings of relevant research, in order to outline more detailed hypotheses for the subsequent analysis.

4.2 Research context

As presented in chapter 2.2.3, research investigating the impact of OI on the recognition of newly-learned L2 words has revealed evidence that written input can facilitate the encoding of difficult L2 phonological contrasts, including unfamiliar orthographic information (Escudero et al., 2008; Showalter & Hayes-Harb, 2013). However, the facilitative effects of written input appear most commonly in cases where written input is both familiar and grapheme-phoneme correspondences (GPCs) are congruent between the L1 and L2. The relevance of perceptual salience has also been demonstrated by research indicating that written input is at best a confirmatory cue for what can already be perceived, implying limited use for confusable L2 contrasts (Escudero, 2015; Rafat, 2016). Furthermore, several studies have emphasised potential interference from familiar but incongruent OI, resulting in cross-linguistic transfer effects and errors (Bürki et al., 2019; Cerni et al., 2019; Erdener & Burnham, 2005; Escudero et al., 2014; Hayes-Harb et al., 2010; Rafat & Stevenson, 2018).

Notably, the above studies focus on comparisons of European languages which share the Roman alphabet. However, research investigating orthography and L2 phonology across different writing systems is growing. Much of the research to date has focused on English naïve learners of pseudowords with entirely unfamiliar written input (Hayes-Harb & Cheng, 2016; Jackson, 2016; Mathieu, 2016; Shepperd, 2018; Showalter, 2018; Showalter & Hayes-Harb, 2015). These studies demonstrate minimal influence of novel graphemes when lexically encoding confusable L2 contrasts, although there are mixed interpretations of positive and negative effects. However, there is persistent evidence that distinct script input is less disruptive than familiar incongruent OI (Hayes-Harb & Cheng, 2016; Jackson, 2016; Showalter, 2018). Across both shared and distinct script input, the precision of phonological representation of the target contrasts is a contributing factor (Mathieu, 2016; Ota et al., 2009; Shepperd, 2013, 2018), as is L2 proficiency (Escudero et al., 2014; Hao & Yang, 2021; Mok et al., 2018; Showalter, 2020; Veivo et al., 2016; Veivo & Jarvikivi, 2013). For example, the potential for written input to interfere with lexical encoding increases if a phonological contrast is not well-established and if L2 proficiency is lower.

Overall, these studies demonstrate that orthographic systematicity, familiarity, congruence, perceptibility and L2 proficiency are important factors to consider when investigating the effects of orthographic input on L2 phonological acquisition. Several of these studies implement similar measures of lexically encoding confusable contrasts as found in the present study, namely a word learning phase followed by an audio-visual matching task. Even though sample sizes were often small (n = ~15-30), they offer a basis from which to make predictions for the present study. The only notable change is that of terminology, where *shared* vs. *distinct* is used, rather than *familiar* vs. *unfamiliar*, as this separates orthographic knowledge of L1 and L2 scripts from proficiency. In particular, as all L1 Arabic participants are literate in the English script, the term *unfamiliar* does not capture the relationship between the input and the learners' literacy experience. Thus, the following predictions are made for the analysis presented in this chapter:

1) **Perceptibility**: L1 Arabic-speakers will have more difficulty lexically encoding the contrast between words that differ by /f-v/ compared to /m-n/, as it is a nonnative

contrast that typically assimilates to a single /f/ sound in Arabic. This will be evidenced in lower accuracy and slower reaction times when recognising newly-learned minimally distinct words differing by /f-v/ compared to /m-n/ across all written input conditions.

2) **Incongruent shared OI**: Exposure to Arabic OI during word learning will lead to interference effects for /f-v/ words, as the incongruent shared transliteration of the two sounds with a single letter in Arabic will encourage spurious homophony. This will be evidenced in lower accuracy and slower reaction times compared to both English OI and no OI conditions for words differing by the target contrast.
3) **Systematic distinct OI**: Exposure to English OI during word learning will lead to facilitative effects for /f-v/ words, as the systematic representation of the two sounds with two separate letters will visually disambiguate the speech signal. This will be evidenced in higher accuracy and faster reaction times compared to both Arabic OI and no OI conditions for words differing by the target contrast.
4) **L2 Proficiency**: The influence of orthography will be mediated by L2 English proficiency, where increasing familiarity with the script and target language phonology will be evident in higher accuracy and faster reaction times for both English and Arabic OI compared to no OI. Performance will improve with English OI as familiarity improves rapid integration of congruent audio and visual information in memory. Meanwhile, performance will improve with Arabic OI as more experienced learners are better able to overcome interference effects.

4.3 Methodology

To reiterate the methods applied to address this question, participants were required to learn 12 minimally distinct English pseudowords, differing either by /m-n/ or by /f-v/, and an accompanying image. These words were presented with the Arabic spelling, English spelling or without any written form. Participants saw each of the target items four times in a random order. Words were presented in pairs and after each pair they were asked to listen and identify the words again from a grid containing the corresponding images and two distractors. The pairs were always from the same orthographic condition but never minimal pair items. Examples of presentation and testing trials for the words learning phase are provided in figure 4.1 and 4.2.

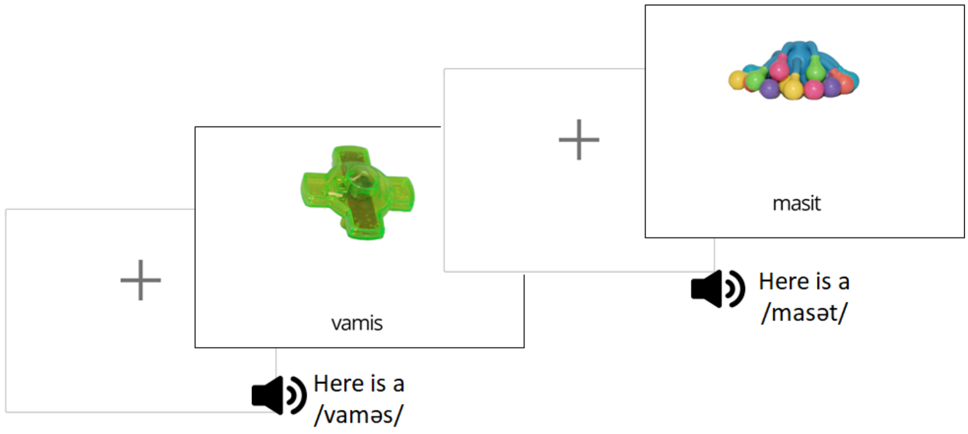


Figure 4.1: Word learning phase presentation trial sequence

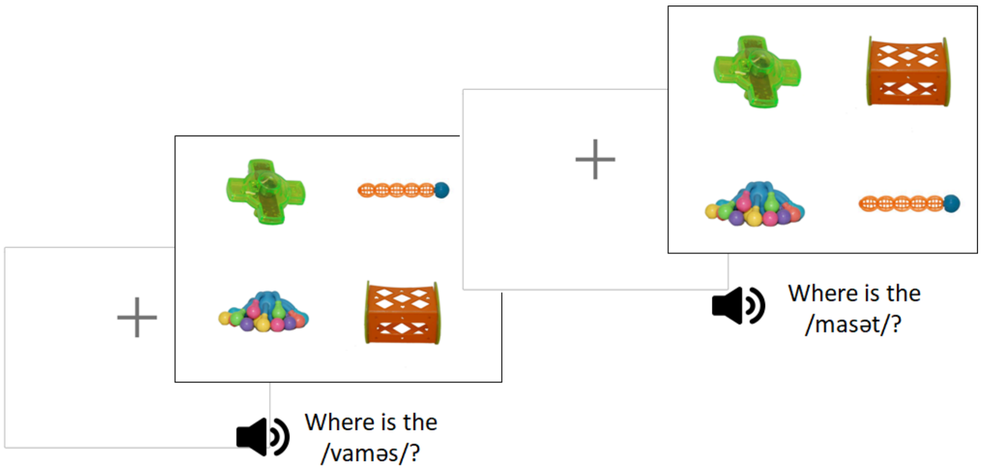


Figure 4.2: Word learning phase testing trial sequence

After the learning phase, participants saw the same images and heard the same words, without any written forms, and had to decide whether the word and image matched. Half of the trials matched and the other half that did not match presented the images of minimal pair items (figure 4.3). Participants saw each of the items twice in both match and mismatch trials, meaning there were 48 trials for each participant. It was assumed that participants who were unable to correctly identify when the image and audio match and mismatch had not lexically encoded the target phonological contrast in sufficient detail; thus, encoding the minimal pair items as homophonous.

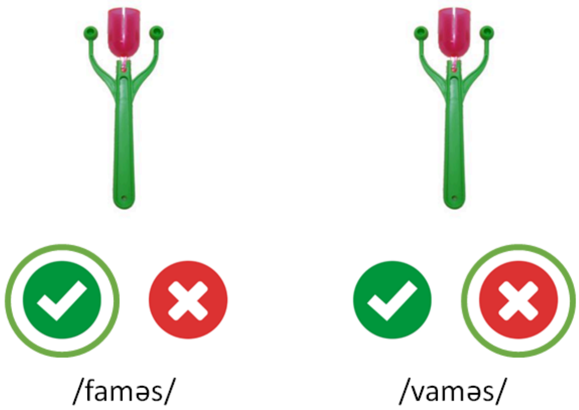


Figure 4.3: Example match and mismatch trial screens

The data from the audio-visual matching task was merged with relevant information from the audio check, demographic questionnaire, word learning task and distraction check. An overview of how the data was processed and explored before the main analysis can be found in appendix V and VIII. The analysis presented here is based on the data from 118 L1 English-speakers and 114 L1 Arabic-speakers, as outlined in 3.3.2. The following sections outline the analysis of participant accuracy and response times for the word learning and word matching stages of the study. The purpose of analysing the results from the word learning phase is to assess the extent to which participants were able to learn and recognise novel items after a short period of training, as well as check for particular difficulty with items differing by each phonological contrast and presented in each OI condition.

Descriptive overviews of mean correct responses and response times for both language groups are provided, followed by further analysis using generalised linear mixed models (GLMMs). Both correct responses and response times for the word learning phase are analysed with GLMMs using the lme4 package in R. Analysis using GLMMs is also applied to the accuracy data from the matching task. However, for the main analysis of the matching task, the data from the two language groups is analysed separately. This is because the L1 English group was included as a control group, to ensure the validity of the test and not as a native-like target for the L1 Arabic group. Data for response times is descriptively reported for additional insight. However, due to the length of time taken to respond to items, it was clear that these times did not reflect unconscious processes or rapid reactions, but rather longer conscious processes to recall items. Thus, offering less insight into the mental representations of novel lexical items.

There are several advantages to using mixed-effects models to analyse the data from these tasks, such as the opportunity to take variability within and across both participants and items into account. In contrast to ANOVAs, observations are not aggregated across participants and coefficient estimates of fitted models indicate both the direction and magnitude of effects. Additionally, mixed models can better deal with unbalanced designs and both continuous and categorical data, offering a more flexible tool (Baayen et al., 2008; V. A. Brown, 2021; DeBruine & Barr, 2019). The R Markdown (Rmd) files with the full analysis scripts can be found in appendices IX and X.

4.4 Word learning results

The mean accuracy and response times, summarised in table 4.1 and figure 4.4, show L1 Arabic-speakers performed with lower accuracy and slower response times than L1 English-speakers. However, both groups demonstrated mean performance of above 90% accuracy and ceiling effects. Thus, there was no indication of particular difficulty with the initial task of memorising the words and distinguishing them from a non-minimal pair item or distractor images. Due to the length of time taken to respond, timings are not referred to as 'reaction times', but rather as 'response times', as these times are more likely to measure the time it takes for participants to consciously recall and respond to test items.

Table 4.1: Mean accuracy and response times from word learning phase, by L1 groups

	L1 Arabic	**L1 English**
Mean proportion correct (*SD*)	.92 (.27)	.95 (.22)
Mean response time (ms) (*SD*)	2075 (1321)	1949 (1039)

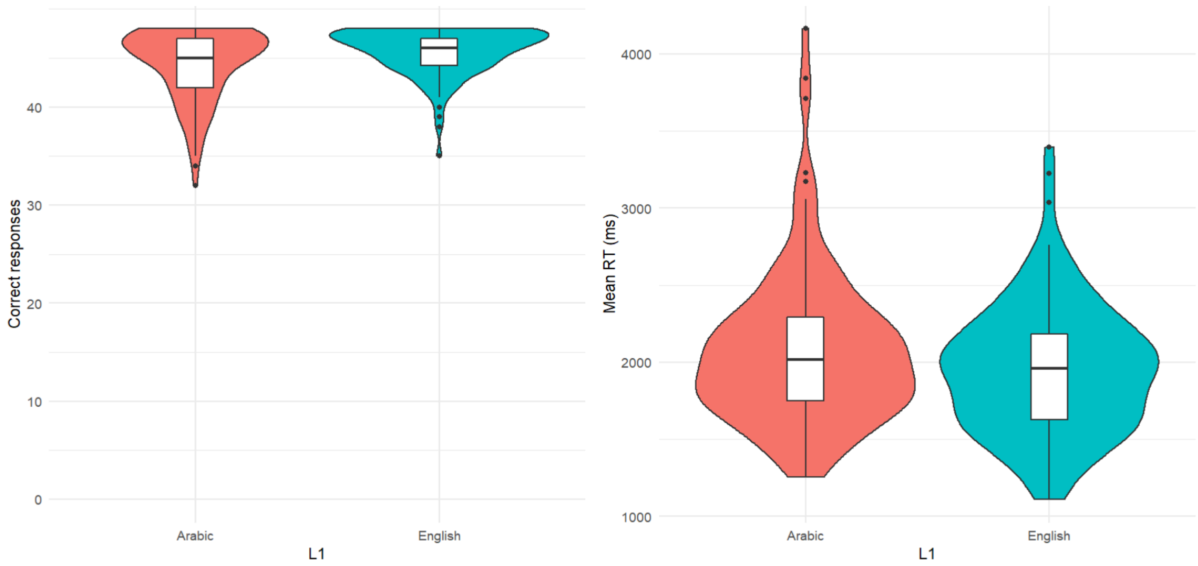


Figure 4.4: Word learning raw accuracy and mean response time data

4.4.1 GLMMs with word learning accuracy

Accuracy was then analysed using mixed-effects modelling to assess whether performance was significantly different between the two groups, as well as exploring the influence of phonological contrast and OI. A stepwise procedure was followed to build the model, whereby the theoretically motivated fixed effects of L1 group, orthographic input and phonological contrast were initially added to the model. Fixed effects of L1, OI and phonological contrast were contrast coded to centre the variables and aid interpretation of the model. The two-level variables of L1 group and phonological contrasts were sum coded (*L1 Arabic* 1, *L1 English* -1; *fv* 1, *mn* -1). Meanwhile, the three-level factor of OI was helmert coded, to facilitate the comparison between any OI and no OI, then Arabic OI and English OI (Brehm & Alday, 2022; Shad et al., 2020). Additionally, a maximal random effects structure was adopted, which included random effects with varying intercepts by participant and by item. The maximal model also included random slopes for OI and phonological contrast by participant. However, these slopes were not included in the final model due to convergence problems that arose when modelling the data, as well as avoiding loss of power through overly complex random effects (Matusheck et al., 2017).

Model comparisons with likelihood tests explored interactions and improved fit with additional fixed effects, based on demographic and environmental variables. Such variables included: age, level of completed education, audio setup, device type and amount of distraction. Likelihood ratios were calculated using the anova() function in R and compared the full model to a model without each fixed effect and interaction (Winter, 2019). The chi-square results reported alongside model estimates below indicate whether effects and interactions significantly contributed to the final model. The final model is reported below:

glmer (Correct ~ L1 + OI + contrast + age + (1 | participant) + (1 | item),

data = learn_data2, family = "binomial",

Control = glmerControl (optimizer = "bobyqa", optCtrl=list(maxfun=100000)))

This analysis revealed a main effect of L1, where accuracy was lower in the L1 Arabic group than in the L1 English group (β = -0.29, SE = 0.08, $\chi^2(1)=14.4$, p<0.001). There was also a main effect of OI, where accuracy was lower when words were accompanied with any OI compared to the audio only condition (β = -0.27, SE = 0.13, z(11140) = -2.07, p=.04); however, likelihood testing revealed that inclusion of this factor as a fixed effect did not significantly improve the model fit ($\chi^2(2)=3.67$, p=0.15). Words differing by the two contrasts or taught with English OI compared to Arabic OI did not significantly predict accuracy. There were also no significant interactions between these fixed effects. However, there was an additional main effect of age, where accuracy was lower when participants were older (β = -0.02, SE = 0.01, $\chi^2(1)=13.8$, p<0.001). Table 4.2 provides a model summary for word learning accuracy.

Table 4.2: GLMM model summary for word learning accuracy

Predictors	**Odds ratios**	**CI**	***p***
(Intercept)	53.57	32.32 – 88.80	**<0.001**
L1 (Arabic)	0.75	0.64 – 0.86	**<0.001**
OI vs. no OI	0.76	0.59 – 0.99	**0.039**
English OI vs. Arabic OI	0.99	0.74 – 1.32	0.940
phonological contrast (f-v)	0.97	0.86 – 1.10	0.677
age	0.97	0.96 – 0.99	**<0.001**
Random Effects			
σ^2	3.29		
$\tau_{00\ Participant}$	0.76		
$\tau_{00\ Item}$	0.03		
ICC	0.19		
$N_{Participant}$	232		

N_{Item}	12
Observations	11140
Marginal R^2 / Conditional R^2	0.035 / 0.222

4.4.2 GLMMs with word learning response times

Reaction time data is typically positively skewed and not normally distributed, which violates the normality assumptions of linear mixed models (LMMs). This was the case for the response time (RT) data in this study (figure 4.5). While it is common to use transformations to normalise the distribution, Lo and Andrews (2015) highlight the drawbacks to this approach. They outline the advantages of GLMMs assuming an Inverse Gaussian distribution and using the identity link function in R for psycholinguistic RT data. The present analysis aligns with their proposals to avoid distorting the ratio scale of raw RT data, under the assumption that differences in RT directly reflect the amount of time taken to perform these mental operations, according to mental chronometry approaches. Thus, to avoid the loss of differences between the two language groups through a log-transformed scale, the GLMM approach was preferable. However, as mentioned, the long response times here are not indicators of automatic reactions. It is also impossible to know whether longer responses reflect thoughtful consideration or lack of attention.

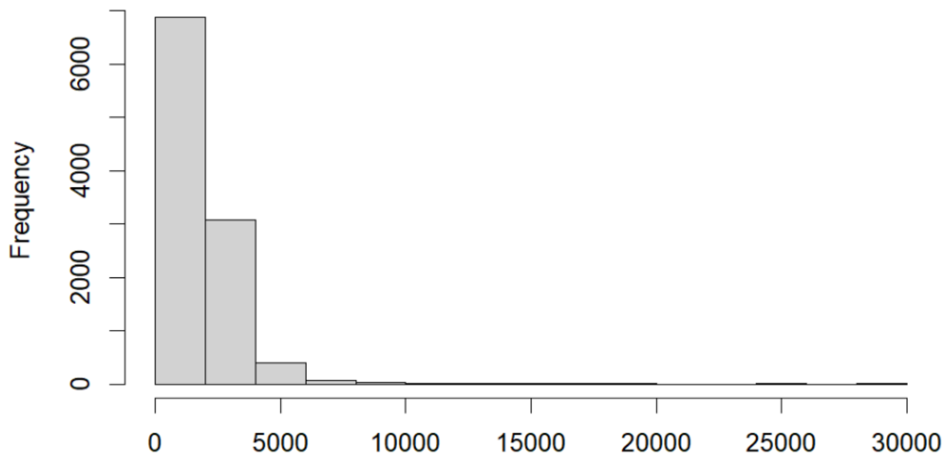


Figure 4.5: Histogram of response time data for the word learning phase

A GLMM was run with response time data, where only times for correct responses were included in the analysis. A similar stepwise procedure was followed to build the model as described for the accuracy data. Additionally, device type was included in the RT model.

This three-level variable was sum coded for comparisons of RT with a computer and mobile phone to the general mean, which included the third level of tablet response times. As above, L1 and phonological contrast levels were also sum coded. However, in order to fit the model using the Inverse Gaussian distribution, levels were coded in the opposite direction, so that the levels with lower predicted performance were coded as the reference levels (i.e., *L1 English* 1, *L1 Arabic* -1; *mn* 1, *fv* -1). As before, a maximal random effects structure was adopted. However, due to issues around singularity of fit, OI was dropped as a random slope by participants. Again, model comparisons with likelihood tests explored interactions and improved fit with additional fixed effects based on demographic and environmental variables. While the model was reported to have converged using the bobyqa optimiser, inspection of the random effects summary in table 4.3 reveals issues with the model fit in the random effects correlation, and in turn the conditional R^2. Therefore, the results below are interpreted with caution. The final model and summary are reported below:

glmer (Reaction.Time ~ L1 + OI + Contrast + Participant.Device.Type +
(1 + Contrast | Participant) + (1 | Item),
data = right_learn_data2, family = inverse.gaussian (link ="identity"),
control = glmerControl(optimizer = "bobyqa", optCtrl = list(maxfun=100000)))

This analysis revealed a main effect of L1, where responses were faster in the L1 English group than in the L1 Arabic group (β = -68.74, SE =13.82, $\chi^2(1)$=18.35, $p<0.001$). There was also a main effect of OI, where responses were slower when words were presented with OI compared to the audio only condition (β = 45.36, SE =15.65, t(10447) = 2.90, p=0.004); however, likelihood testing revealed that inclusion of this factor as a fixed effect did not significantly improved the model fit ($\chi^2(2)$=1.38, p=0.5). There was also a main effect of contrast, where participants responded more quickly to words differing by /m-n/ than by /f-v/ (β = -31.45, SE =11.57, $\chi^2(1)$=16.01, $p<0.001$). In addition, there was a main effect of device type, where response times were slower when participants used a computer (β =177.73, SE = 12.92, t(10447) = 13.76, $p<0.001$); however, the contribution of this effect to the model was only approaching significance ($\chi^2(2)$=4.97, p=0.08). Words taught with English OI compared to Arabic OI did not significantly predict response times and there were no significant interactions between these fixed effects.

Table 4.3: GLMM model summary for word learning response times

Predictors	**Estimates**	**CI**	***p***
(Intercept)	2075.79	2048.69 – 2102.88	**<0.001**
L1(English)	-68.74	-94.56 – -42.92	**<0.001**
OI vs. no OI	45.36	14.70 – 76.03	**0.004**
English OI vs. Arabic OI	-21.46	-44.48 – 1.57	0.068
Contrast (m-n)	-31.45	-54.13 – -8.76	**0.007**
Participant Device (Computer)	177.73	152.41 – 203.04	**<0.001**
Participant Device (Mobile)	19.06	-7.41 – 45.53	0.158
Random Effects			
σ^2	0.01		
$\tau_{00\ ID}$	75283.34		
$\tau_{00\ Item}$	1391.91		
$\tau_{11\ ID.Contrast}$	14113.29		
$\rho_{01\ ID}$	-0.04		
ICC	1.00		
N_{ID}	232		
N_{Item}	12		
Observations	10447		
Marginal R^2 / Conditional R^2	0.104 / 1.000		

4.4.3 Word learning summary

Both groups completed the word learning task with ceiling levels of accuracy, with mean scores over 90%. This means that after exposure to target items, followed by a short interval, participants were sufficiently attentive to establish the words and images in working memory and then recognise them with a high level of accuracy from non-minimal pair words and distractor images. Additionally, this demonstrates that even with the high level of novelty, all participants were able to complete this phase of the study with relative ease. Some participants reported using different memory techniques to focus on specific differences between the pairs that were taught together rather than trying to learn the word forms. However, it appears that overall participants were actively engaged in the word learning phase.

Further analysis of the accuracy and response times during word learning revealed that the L1 English group significantly outperformed the L1 Arabic group, both in accuracy and response times. This is unsurprising as task demands were higher for the L1 Arabic group, due to the (1) L2 English-like quality of pseudowords, including nonnative phonological contrasts, (2) embedding of words in L2 English carrier phrases, (3) switching demands between L1 Arabic instruction screens and L2 English trials, and (4) switching demands between Arabic and English orthographies, as the L1 Arabic group were literate in both. There was no significant effect of phonological contrast with regards to accuracy, although recall demands appeared to be higher for the words differing by /f-v/ compared to /m-m/, demonstrated in the significantly longer response times. Additionally, while OI was a significant main effect for both models of accuracy and response time, likelihood ratios revealed that in both cases inclusion of OI as a fixed effect did not significantly improve the fit of the model. Thus, further interpretation is not well supported by the data.

Additional effects were found for both age and device type, where older participants responded with lower accuracy and those using a computer exhibited slower responses. The age effects are likely explained by the working memory and auditory demands of the task, which would therefore become more difficult as working memory capacity and hearing ability declines with age. Finally, slower responses with the use of computers, but not mobile phones is most likely explained by the difference between recording a touchscreen response and using mouse/mousepad to record a response. Importantly, device type did not predict accuracy during word learning. In sum, all participants were able to complete the word learning phase with a high degree of accuracy across all items and experimental conditions. The next section turns to performance on the audio-visual matching task to investigate whether participants were able to lexically encode the words in memory with sufficient detail to distinguish between novel minimal pair items.

4.5 Word matching results

4.5.1 Descriptive overview

Participant responses in the matching task were binary coded for accuracy with levels ‘correct’, where either a match or mismatch was correctly identified, and ‘incorrect’ where a match was identified as a mismatch or vice versa. Table 4.4 offers a summary of mean percentage correct responses for the language groups, alongside d-prime (*d'*) and criterion scores (*c*), calculated according to signal detection theory (Stanislaw & Todorov, 1999). Signal detection theory addresses the possibility of response bias in the yes-no task and assesses the sensitivity to 'noise' (mismatch) vs. 'signal' (match) trials. D-prime scores were calculated as $d' = z(\text{hit rate}) - z(\text{false alarm rate})$. The hit rate and false alarm rate were adjusted with the log linear approach, to avoid extreme values of 0 or 1 (Hautus, 1995). In addition, criterion scores assess the direction of bias and calculated as $c = ((z\ (\text{hit rate}) + z\ (\text{false alarm rate})) / 2) * 1)$ (Stanislaw & Todorov, 1999). According to this calculation, $c = 0$ would indicate a lack of bias, while a positive *c* value indicates a more relaxed criterion for responding "yes" and a negative *c* value indicates more caution to responding "yes" when discriminating matches. While correct responses indicate the accuracy of recall (i.e., the ability to identify all correct items), *d'* offers more insight into precision (i.e., the ability to reject all incorrect items). There can be a trade-off between recall and precision, but both are required for high performance phonological processing (Pierrehumbert, 2016).

Table 4.4 illustrates that L1 English-speakers outperformed L1 Arabic-speakers in terms of percentage correct answers, as well as higher sensitivity to signal and noise trials. Both groups demonstrate a 'yes' bias, where words and images were more likely to be accepted as a match than rejected. However, the bias was stronger with L1 Arabic-speakers, offering preliminary evidence that minimal pair items were incorrectly accepted as homophonous more so than the L1 English-speakers. This also demonstrates the value of looking at participant mismatch trial performance in particular, as the predicted orthographic and phonological effects are likely to be most evident with the increased difficulty of rejecting a mismatch than accepting a match.

Table 4.4: Summary of mean % correct responses in match and mismatch trials, *d'* and *c* scores from the matching task

	% Correct			
	Match	**Mismatch**	***d'* (*SD*)**	***c* score (*SD*)**
L1 Arabic (*n*=114)	83.6%	49.8%	0.83 (1.10)	0.42 (0.47)
L1 English (*n*=118)	89.3%	67.5%	1.42 (1.02)	0.27 (0.41)

The influence of phonological contrast and orthographic input on correct responses by the two language groups is summarised in tables 4.5 and 4.6. While the accuracy of the L1 English group was higher across both phonological contrasts, both groups found /f-v/ words more difficult than /m-n/ words. This was anticipated for the L1 Arabic group but less so for the L1 English group and is further discussed in section 4.5.2. L1 English participants were still able to respond to /f-v/ items with above chance accuracy, which was not the case with the L1 Arabic group, who exhibited particularly low *d'* and high *c* scores. This indicates that L1 Arabic participants had not lexically encoded words differing by /f-v/ with sufficient detail to distinguish between minimal pair items. The accuracy of the L1 English group also remained higher across all orthographic input conditions compared to the L1 Arabic group. Both groups performed best with English spelling and worse with Arabic spelling. However, the difference between each OI condition in the different language groups was minimal in comparison to the effect of phonological contrast. The relationship between these two variables and raw accuracy are visualised in figure 4.6.

Table 4.5: Summary of mean % correct responses, *d'* and *c* scores from the matching task by phonological contrast

	Contrast	**% Correct**	***d'* (*SD*)**	***c* (*SD*)**
L1 Arabic	/f-v/	60.8%	0.53 (1.03)	0.48 (0.49)
	/m-n/	72.7%	1.13 (1.08)	0.35 (0.44)
L1 English	/f-v/	72.8%	1.14 (1.02)	0.35 (0.45)
	/m-n/	83.9%	1.70 (0.94)	0.19 (0.36)

Table 4.6: Summary of mean % correct responses, *d'* and *c* scores from the matching task by OI condition

	OI	**% Correct**	***d'* (*SD*)**	***c* (*SD*)**
L1 Arabic	none	67.2%	0.85 (1.02)	0.41 (0.47)
	Arabic	63.6%	0.68 (1.09)	0.46 (0.47
	English	69.4%	0.97 (1.16)	0.38 (0.46)
L1 English	none	79%	1.45 (1.00)	0.31 (0.41)
	Arabic	76.5%	1.32 (1.05)	0.27 (0.45)
	English	79.7%	1.49 (1.00)	0.23 (0.38)

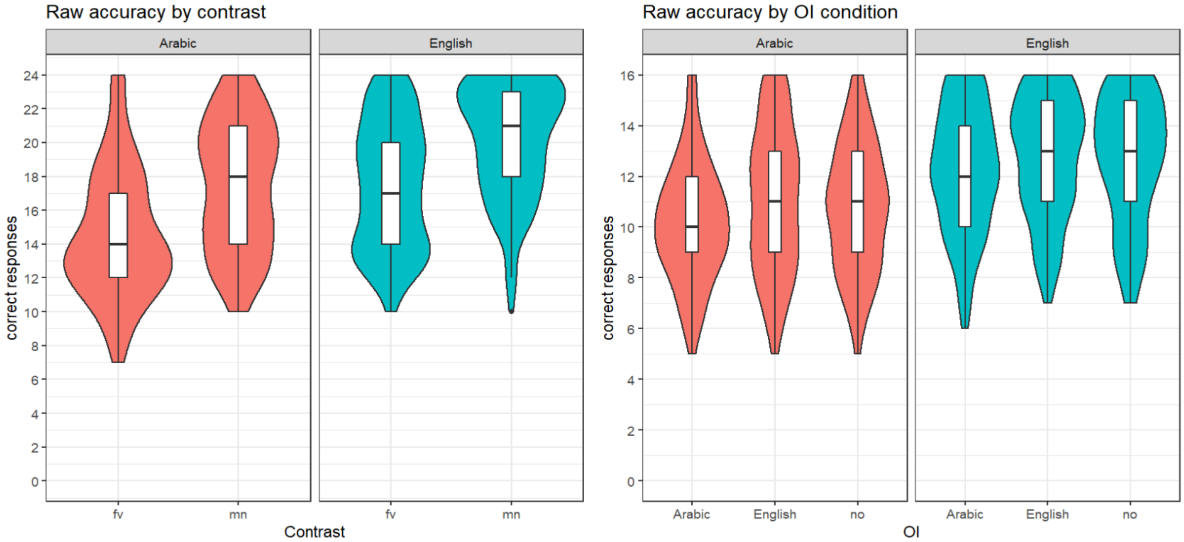


Figure 4.6: Matching task correct answers for each L1 group by contrast and OI condition

Mean participant response times were calculated, based on correct responses only, and are summarised for the two language groups in table 4.7. The response times mirror the accuracy result above, as L1 English-speakers also outperformed L1 Arabic-speakers with faster responses. As mentioned with the word learning phase, the length of time taken to respond to items was far longer than what would be expected for an automatic reaction and is assumed to reflect participants' conscious processing and recall of items. It was noted that longer times could be related to lack of attention or reduced pressure to complete the task in a timely manner outside of the lab. The lack of direct contact with the participants means it is difficult, if not impossible, to assess the extent to which this may have been the case. Participants were told that their place on the leaderboard was based on both their accuracy and reaction time, with the aim that they would be incentivised to attend and respond both quickly and accurately. While feedback from participants indicated that they were incentivised by the leaderboard, it also appeared to be the case that some participants who perceived themselves to be performing poorly lost motivation to fully attend. In these instances, participants were more likely to rush through trials, suggesting that longer times reflected thoughtful recall more than inattention. This also shows the reasoning behind only exploring the response times for correct answers. The amount of speculation surrounding the interpretation of response times means that only a descriptive overview of response times is offered in relation to the matching task. Additionally, the unequal loss of observations for the L1 Arabic vs. L1 English group, and the /f-v/ contrast words in comparison to /m-n/ words, from excluding incorrect responses did not form a good basis for further statistical analysis. The extreme values visible in figure 4.7 may have also contributed to difficulty fitting models to the data.

Table 4.7: Mean response times from the matching task for each L1 group

	L1 Arabic	**L1 English**
Mean response time (ms) (*SD*)	2235 (2623)	1917 (1303)

The influence of phonological contrast and orthographic input on response times by the two language groups is summarised in table 4.8. Again, mirroring the accuracy results, the L1 English group was faster across both contrasts and all OI conditions, compared to the L1 Arabic group. Both groups were slower at responding to items differing by /f-v/ than /m-n/, with more variation around the mean. In addition, they exhibited a similar pattern of response times across the different OI conditions, where mean response time was fastest in the audio-only condition and slowest when words were accompanied by Arabic OI. This differs slightly from the accuracy results, where raw accuracy appeared marginally higher with English OI than no OI. It is noteworthy that variance around the mean was far higher for the audio-only words for the L1 Arabic group in particular, demonstrating there was not necessarily a clear advantage associated with a lack of orthographic input. In contrast, the slower mean response times and greater variance for both groups with the Arabic OI imply a broadly inhibitory influence of this condition, especially when taken together with the mean accuracy scores.

Table 4.8: Mean RT (ms) from the matching task for each L1 group by phonological contrast and OI condition (SDs)

	L1 Arabic	**L1 English**
Contrast		
/f-v/	2208 (3374)	1876 (962)
/m-n/	2189 (1803)	1735 (869)

OI		
Audio only	2150 (3767)	1738 (945)
Arabic OI	2240 (1858)	1836 (1094)
English OI	2206 (1773)	1771 (921)

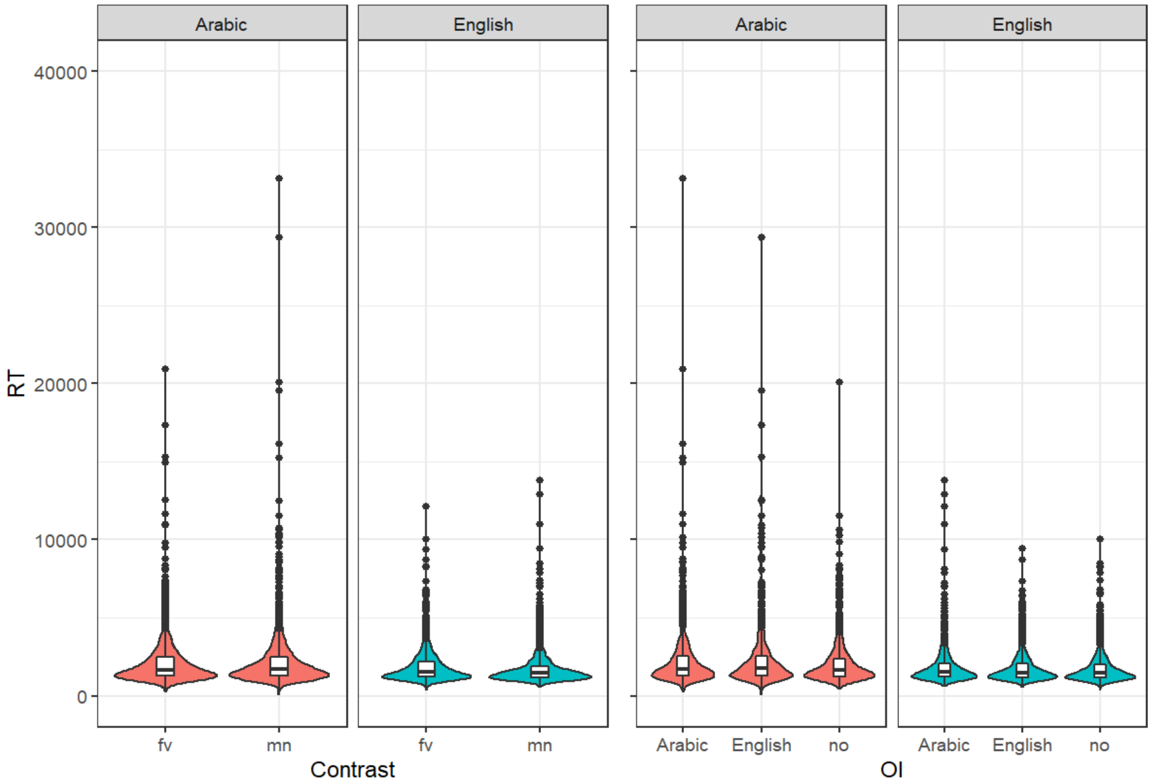


Figure 4.7: Matching task raw RT by contrast and OI condition for each L1 group

At this point it is useful to outline the nature of the comparison between the L1 Arabic-speakers and the L1 English-speakers. It was not the goal of this study to use the L1 English-speakers as a measure of 'native-like' performance by the L1 Arabic learners of English. Rather, they functioned as a control group to assess the validity of the test and aid interpretation of the results of the L1 Arabic-speakers. Therefore, based on the performance of the L1 English-speakers, we can confidently assume that it was possible to complete the audio-visual matching task with a high level of accuracy, after a short word learning phase introducing the novel words. Indeed, even though performance was lower for both the L1 English group and the L1 Arabic group with the target /f-v/ items, it was possible to perform the matching task with above chance accuracy for these items. The same can be said for the words across all three orthographic conditions. Additionally, the high performance for both groups when distinguishing words differing by the easier /m-n/ contrast indicate that the task itself was not the primary source of difficulty for participants. Therefore, it is assumed that the results from this task provide insight into the influence of orthography on a difficult nonnative contrast for L1 Arabic learners of L2 English. However, the difficulty of the /f-v/ words for L1 English-speakers may indicate an issue with those test items, which is explored in the following section.

4.5.2 GLMMs with L1 English-speaker accuracy

It was anticipated that the L1 English-speaker group would perform similarly with words differing by both phonological contrasts, as the words followed L1 phonotactic constraints and minimal pairs differed by phonological contrasts that were well-established in English. It was also anticipated that orthographic information in Arabic spelling when learning the new words would be inaccessible, as none of the participants report any Arabic literacy ability. This would lead to comparable performance between words taught without any written input, or would create a visual distraction that would reduce accuracy. In comparison, performance with the English spelling was predicted to be higher, especially as both phonological contrasts were assumed to be easily perceivable. As indicated by the descriptive overview of accuracy, these predictions were not entirely borne out. To further assess the claims, the results were analysed using GLMMs run in R, using the lme4 package.

As with the word learning analysis, a stepwise procedure was followed to build the model, whereby the theoretically motivated fixed effects of trial type (match vs. mismatch), orthographic input and phonological contrast were initially added to the model. Fixed effects of trial type, OI and phonological contrast were contrast coded to centre the variables and aid interpretation of the model. The two-level variables of trial type and phonological contrasts were sum coded (*mismatch* 1, *match* -1; *fv* 1, *mn* -1). Meanwhile, the three-level factor of OI was helmert coded to facilitate the comparison between any OI and no OI, then Arabic OI and English OI. As before, a maximal random effects structure was adopted, including random effects with varying intercepts by participant and by item. The maximal model included random slopes for trial type, OI and phonological contrast by participant, as well as trial type by item. However, this model did not converge, leading to a model with reduced complexity in the random effects structure. Model comparisons with likelihood tests explored interactions and improved fit with additional fixed effects based on demographic and environmental variables. The model and summary are reported below:

```
glmer (Correct ~ OI * contrast + match + age +
        (1 + contrast | participant) + (1 + match | item) +
        data = ES_data, family = "binomial",
        Control = glmerControl (optimizer = "bobyqa", optCtrl=list(maxfun=100000)))
```

This analysis revealed a main effect of OI, where accuracy was higher when words had been taught accompanied by English spelling compared to Arabic spelling (β = 0.26, SE = 0.13, $\chi^2(4)$=11.2, p=0.02). There were also main effects of phonological contrast, trial type and age. Accuracy was significantly worse (1) with words differing by /f-v/ (β = -0.47, SE = 0.07, $\chi^2(3)$=23.5, p<0.001), (2) in mismatch trials (β = -0.80, SE = 0.08, $\chi^2(1)$=26.1, p<0.001), and (3) when participants were older (β = -0.02, SE = 0.01, $\chi^2(1)$=9.3, p=0.002). Finally, there was a significant interaction between OI and phonological contrast, where accuracy was lower for /f-v/ words taught with any OI compared to audio only (β = -0.40, SE = 0.11, $\chi^2(2)$=9.2, p=0.01). The summaries provided in table 4.9 and figure 4.8 have converted estimates into odds ratios and provided the 95% confidence intervals for clear interpretation of effect sizes.

Table 4.9: GLMM summary for L1 English matching task accuracy

Predictors	**Odds ratios**	**CI**	***p***
(Intercept)	13.21	7.59 – 22.98	**<0.001**
OI vs. no OI	0.97	0.77 – 1.23	0.825
English OI vs. Arabic OI	1.31	1.02 – 1.68	**0.034**
phonological contrast (f-v)	0.63	0.55 – 0.72	**<0.001**
trial type (mismatch)	0.45	0.38 – 0.53	**<0.001**
age	0.98	0.96 – 0.99	**0.002**
OI vs. no OI * contrast (f-v)	0.67	0.54 – 0.84	**<0.001**
English OI vs. Arabic OI * contrast (f-v)	0.89	0.69 – 1.16	0.397

Random Effects	
σ^2	3.29
$\tau_{00\ \text{Participant}}$	0.87
$\tau_{00\ \text{Item}}$	0.02
$\tau_{11\ \text{Participant.contrast (f-v)}}$	0.16
	0.06

τ_{11} Item.tial type (mismatch)	-0.35
ρ_{01} Participant	-0.35
ρ_{01} Item	0.25
ICC	118
N Participant	12
N Item	
Observations	5664
Marginal R^2 / Conditional R^2	0.183 / 0.389

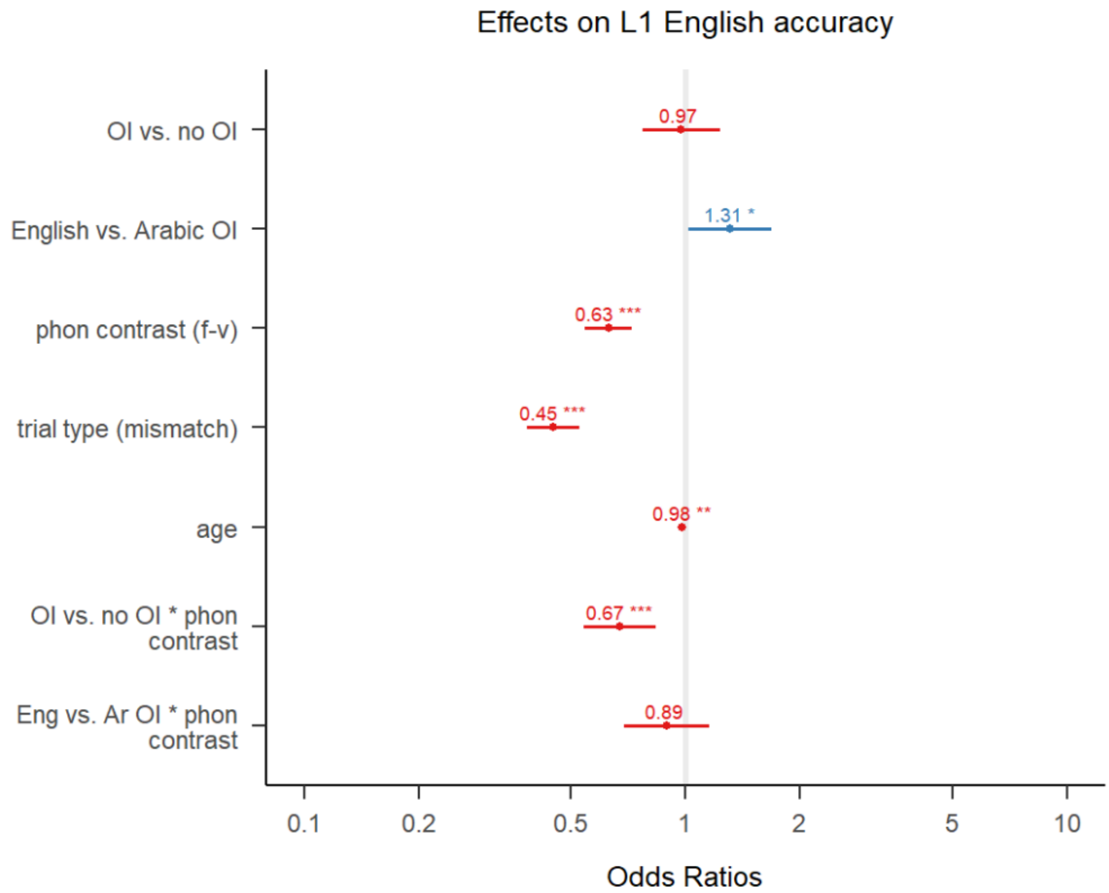


Figure 4.8: L1 English accuracy mixed-model effects, including odds ratios, confidence intervals and level of significance

The largest predictor of participant accuracy was trial type and, due to the anticipated difficulty of the mismatch trials compared to match trials, it was of particular interest how participants performed on those trials. Therefore, for additional insight, the model predictions were plotted for both the main effects of OI condition and phonological contrast, in both trial types separately (figure 4.9). These figures visualise the estimated marginal means of correct answers for L1 English-speakers, demonstrating the slight advantage when words were taught with English spelling compared to Arabic spelling, and the difficulty when words differed by the /f-v/ contrast. Predicted accuracy remains above chance for items differing by /f-v/. However, this could indicate a problem with the items themselves, such as recording quality being insufficient to perceive the voicing distinction, which is further explored below. The interaction between contrast and OI condition shows difficulty was exacerbated by exposure to OI, as accuracy was significantly higher when participants were exposed to the audio only for /f-v/ words (figure 4.10).

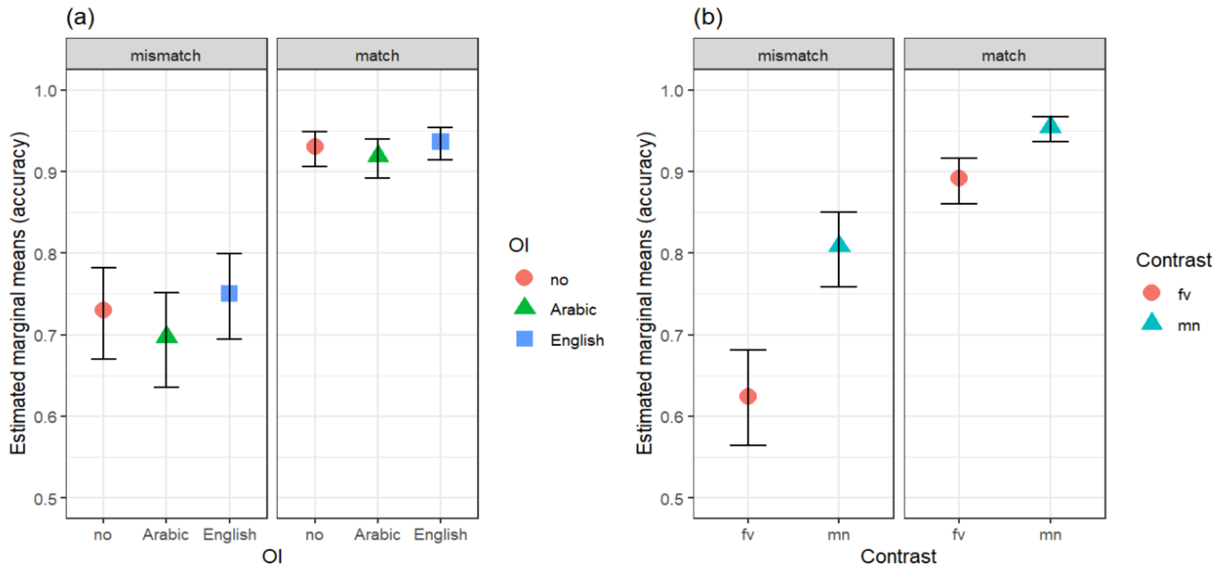


Figure 4.9: Model estimated marginal means for L1 English accuracy, plotted by (a) OI and (b) phonological contrast for both match and mismatch trials

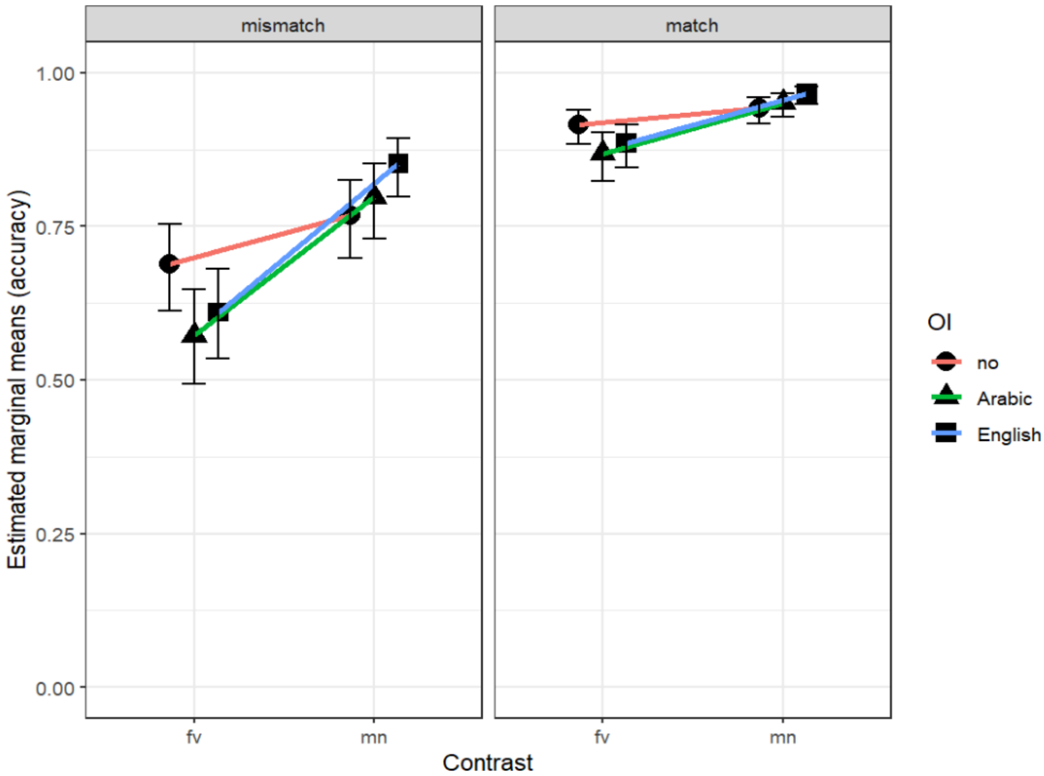


Figure 4.10: Model estimated marginal means for L1 English accuracy, plotting the interaction between phonological contrast and OI for match and mismatch trials

Upon returning to the original audio files, there was evidence of partial devoicing for the words beginning with the voiced labiodental fricatives. Spectrograms of one of the /f-v/ minimal pairs compares the 'voiceless' and 'voiced' counterparts in figure 4.11. Jongman and colleagues (2000) observed noise duration is longer in voiceless compared to voiced fricatives, and can be a salient perceptual cue for a voicing distinction, which is visible in the spectrograms below. However, looking at F0 in the spectrogram for /vædət/, it appears voicing was only partially apparent for the fricative token. Research around the devoicing of fricatives in English suggests that /z/ is more frequently devoiced than /v/ and that devoicing most frequently occurs in the word-final position. However, devoicing and partial devoicing of /v/ is still widely reported (Bjorndahl, 2022). Furthermore, it has been found that the devoicing of /v/ in Standard Southern British English (SSBE) increases for female and older speakers (J. Verhoeven et al., 2011). The audio recordings used in the present study were from a female speaker over the age of 40, who speaks a relatively unmarked variety of British English, which may explain the reduced appearance of voicing for the voiced counterparts of the /f-v/ minimal pair items. Thus, the reduced perceptual cue for the voicing distinction may explain the additional difficulty when completing a task with decontextualised novel words, for both L1 speakers and L2 learners. Meanwhile, the high frequency of /m/-/n/ across the world's languages likely reflects the high acoustic salience of that nasal contrast (Maddieson, 1984; Narayan, 2008; Narayan et al., 2010).

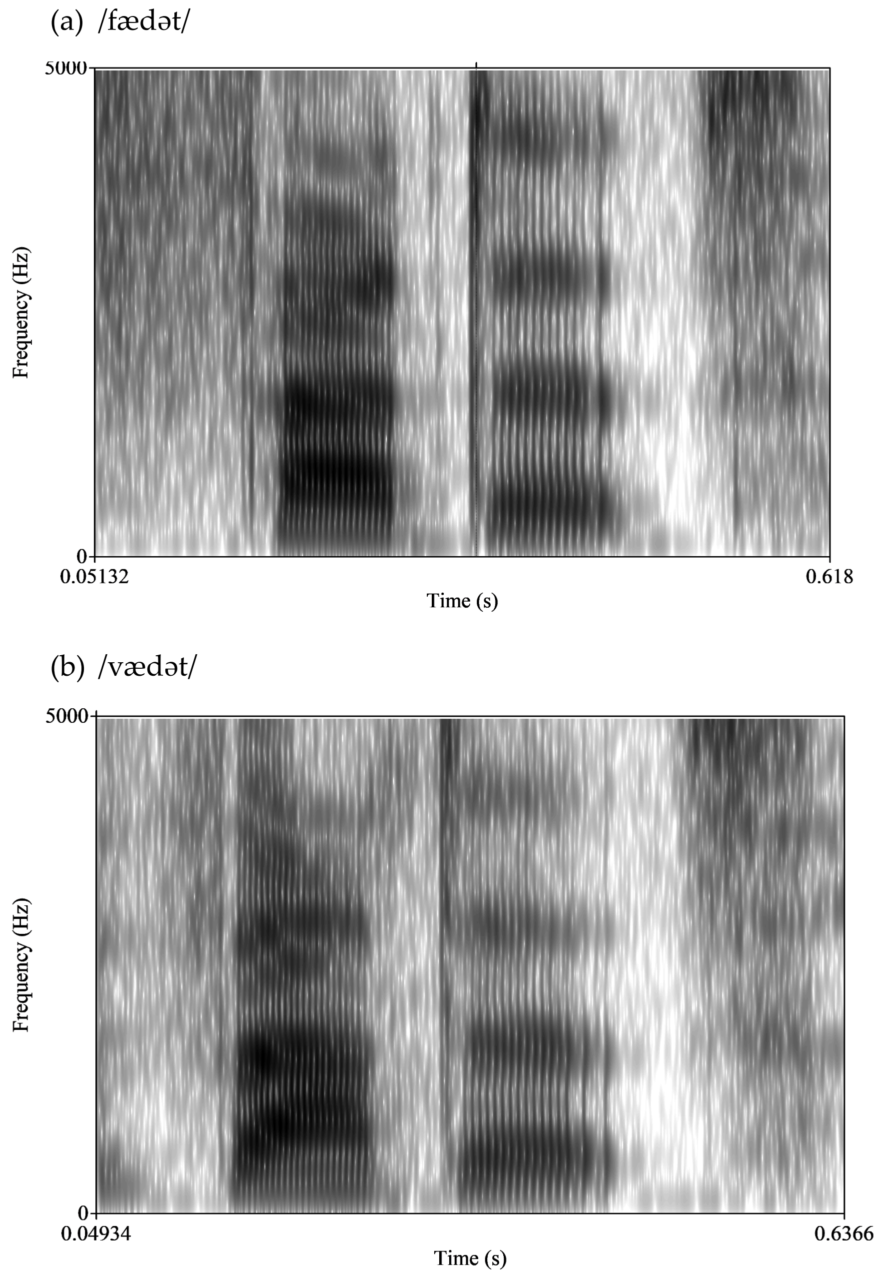


Figure 4.11: Spectrograms of the /f-v/ minimal pair items taught with Arabic OI, as an example of partial word-initial devoicing in (b) /vædət/

4.5.3 GLMMs of L1 Arabic-speaker accuracy

It was predicted that L1 Arabic-speakers would be less accurate when distinguishing /f-v/ minimal pair items than /m-n/ items, especially in light of the low perceptual salience of the contrast demonstrated by the L1 English group. Further to this, it was predicted that this difficulty would be exacerbated by exposure to shared-incongruent Arabic spelling, whereas the English spelling would visually disambiguate the difficult auditory contrast. However, it was anticipated that this would differ depending on L2 English proficiency. To assess these claims, accuracy results were analysed using GLMMs run in R, using the lme4 package.

As with the previous analyses, a stepwise procedure was followed to build the model, whereby the theoretically motivated fixed effects of trial type, orthographic input, phonological contrasts, and proficiency were initially added to the model. As outlined in chapter 3.6, several proficiency measures were included in the demographic questionnaire. These included self-reported level, ability across different skills (reading, writing, listening, speaking), length of study, age of onset, time spent in an Anglophone country, and English language qualifications. To triangulate the self-reported information, a short English proficiency test was administered before participants began word learning. Proficiency measures were explored in terms of their correlations with each other (Appendix V) and correlations with accuracy and response time measures in the matching task (Appendix VIII), leading to the decision to include the score from the English proficiency test in the model as the key measure of proficiency.

As in the previous GLMMs with the L1 English-speaker group, the two-level variables of trial type and phonological contrasts were sum coded (*mismatch* 1, *match* -1; *fv* 1, *mn* -1). Meanwhile, the three-level factor of OI condition was helmert coded to facilitate the comparison between any OI and no OI, then Arabic OI and English OI. The maximal model included random slopes for trial type, OI condition, and phonological contrast by participant, as well as trial type by item. However, this model did not converge, leading to a model with reduced complexity in the random effects structure. Model comparisons with likelihood tests explored interactions and improved fit with additional fixed effects based on demographic and environmental variables. The final model and summary are reported below:

*glmer (Correct ~ OI * contrast + match + proficiency + distraction*

(1 + contrast | participant) + (1 + match | item) +

data = AS_data, family = "binomial",

Control = glmerControl (optimizer = "bobyqa", optCtrl=list(maxfun=100000)))

This analysis revealed main effects of proficiency and orthographic input, where accuracy was higher when participants had higher proficiency in English ($\beta = 0.15$, SE = 0.03, $\chi^2(1)=16.7$, $p<0.001$) and words had been taught accompanied by English spelling compared to Arabic spelling ($\beta = 0.37$, SE = 0.18, $\chi^2(4)=8.2$, $p=0.08$). However, as can be seen from the chi-square result, model comparisons revealed the inclusion of orthographic input as a fixed effect was only approaching significance. There were also main effects of phonological contrast, trial type, and amount of distraction. Accuracy was significantly worse (1) with words differing by /f-v/ ($\beta = -0.37$, SE = 0.09, $\chi^2(3)=13.9$, $p=0.003$), (2) in mismatch trials ($\beta = -0.95$, SE = 0.05, $\chi^2(1)=41.7$, $p<0.001$), and (3) when participants reported higher levels of distraction ($\beta = -0.56$, SE = 0.27, $\chi^2(1)=4.3$, $p=0.04$). Finally, there was a significant interaction between OI and phonological contrast, where accuracy was lower for /f-v/ words taught with any OI compared to no OI ($\beta = -0.47$, SE = 0.18, $\chi^2(2)=6.4$, $p=0.04$). The summaries provided in table 4.10 and figure 4.12 have converted estimates into odds ratios and provided the 95% confidence intervals for clear interpretation of effect sizes.

Table 4.10: GLMM summary for L1 Arabic matching task accuracy

Predictors	**Odds ratios**	**CI**	***p***
(Intercept)	0.73	0.37 – 1.42	0.355
OI vs. no OI	0.99	0.73 – 1.35	0.970
English OI vs. Arabic OI	1.45	1.02 – 2.08	**0.039**
phonological contrast (f-v)	0.69	0.58 – 0.82	**<0.001**
trial type (mismatch)	0.39	0.35 – 0.43	**<0.001**
proficiency	1.16	1.08 – 1.24	**<0.001**
distraction	0.57	0.33 – 0.97	**0.037**
OI vs. no OI * phonological contrast	0.62	0.44 – 0.89	**0.009**
English OI vs. Arabic OI * phonological contrast	0.93	0.59 – 1.46	0.742
Random Effects			
σ^2	3.29		
$\tau_{00\ ID}$	0.50		

$\tau_{00\ \text{Item}}$	0.05
$\tau_{11\ \text{ID.Contrast (f-v)}}$	0.16
$\tau_{11\ \text{Item.trial type (mismatch)}}$	0.01
$\rho_{01\ \text{ID}}$	-0.23
$\rho_{01\ \text{Item}}$	-0.15
ICC	0.18
N_{ID}	114
N_{Item}	12
Observations	5478
Marginal R^2 / Conditional R^2	0.234 / 0.372

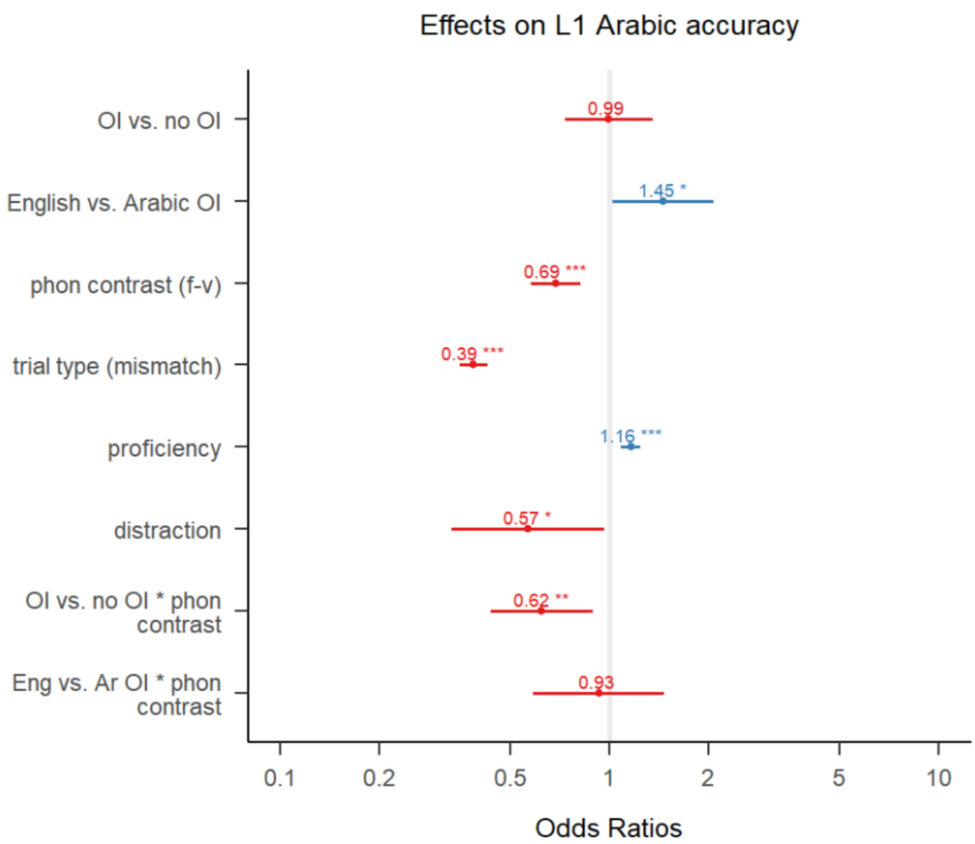


Figure 4.12: L1 Arabic accuracy mixed-model effects, including odds ratios, confidence intervals and level of significance

As with the L1 English group, the strongest predictor of participant accuracy was trial type, which again demonstrates the value of exploring participant performance on the mismatch trials. Therefore, the model predictions were plotted for both the main effects of OI condition and phonological contrast, in each trial type (figure 4.13). Like the L1 English results, these figures visualise the slight advantage when words were taught with English spelling compared to Arabic spelling and the added difficulty when items differed by /f-v/. While the L1 Arabic responses patterned very similarly to the L1 English results, the ability to distinguish between target items in the mismatch trials with above chance accuracy was strikingly distinct. L1 English-speakers' mean accuracy in mismatch trials was above chance across all levels of OI and phonological contrast, whereas L1 Arabic-speakers' mean accuracy in mismatch trials was only above chance when words were learned with English spellings and differed by /m-n/. The difficulty of the /f-v/ words was likely related to the low perceptual salience of a voicing distinction that is not typically present in Arabic varieties, combined with the specific articulation of the target items in the present study. Figure 4.14 then offers a visualisation of the interaction between OI condition and phonological contrast, where it was particularly clear in mismatch trials that accuracy was significantly lower when /f-v/ words were taught with any OI, while the opposite was true for the easier /m-n/ words.

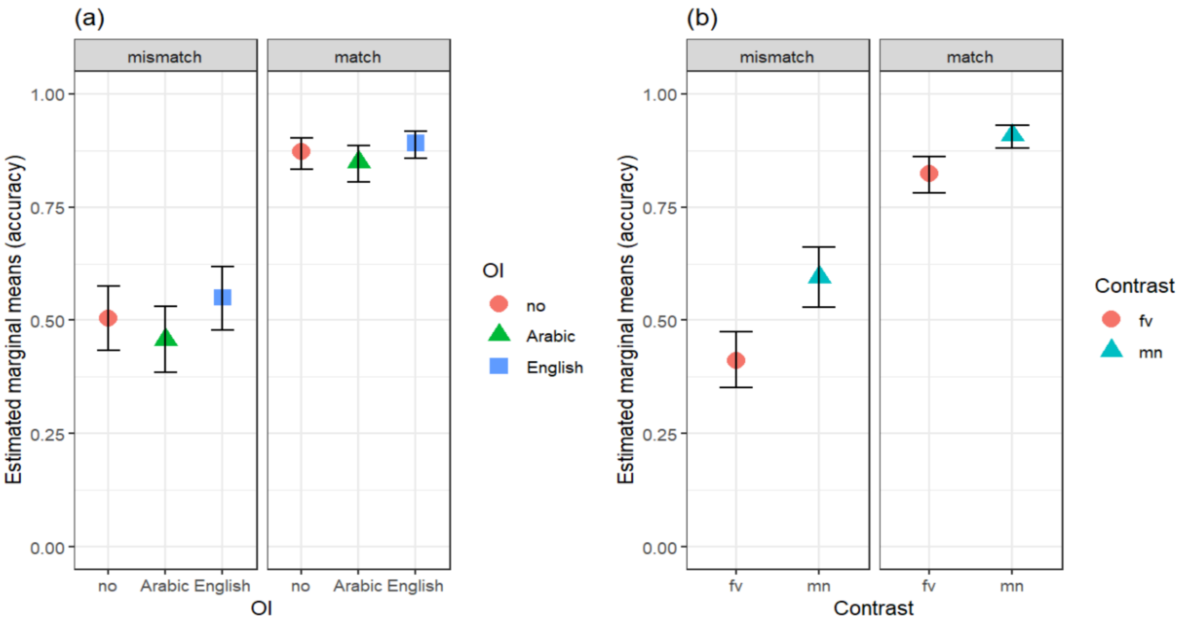


Figure 4.13: Model estimated marginal means for L1 Arabic accuracy, plotted by (a) OI and (b) phonological contrast for both match and mismatch trials

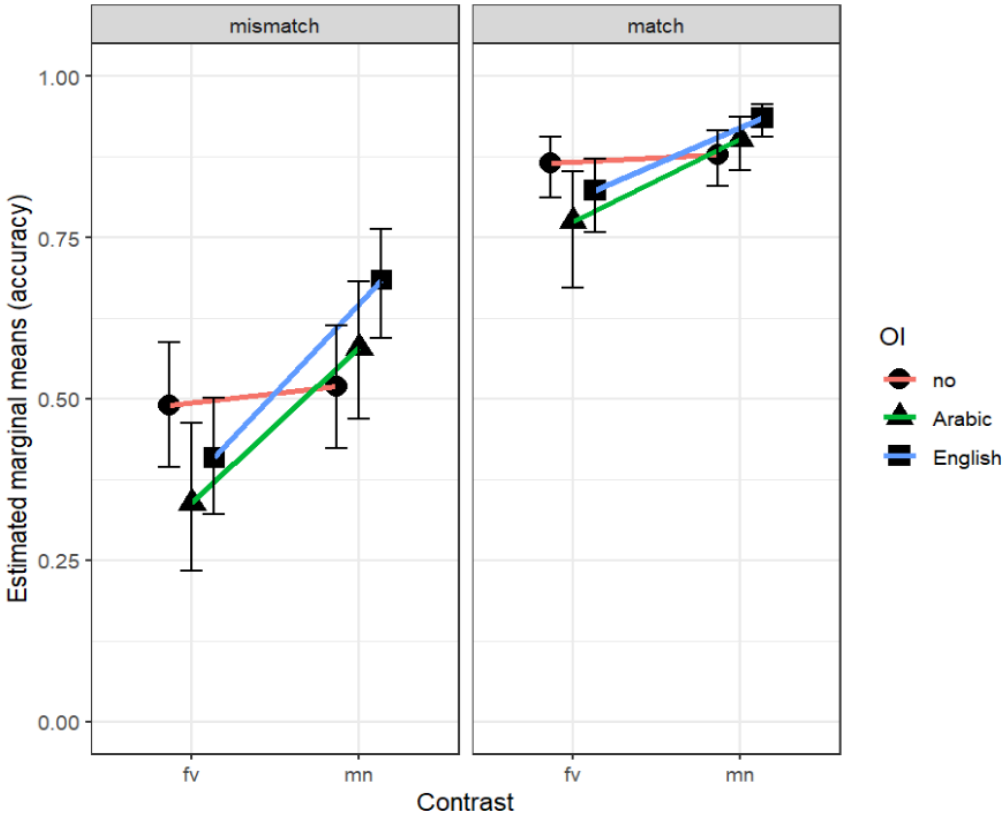


Figure 4.14: Model estimated marginal means for L1 Arabic accuracy, plotting the interaction between phonological contrast and OI for both match and mismatch trials

Another important factor for this analysis was L2 English proficiency, which significantly predicted higher accuracy. Looking at performance in mismatch trials, figure 4.15 demonstrates that, while accuracy improved as proficiency increased, only those with the highest performance on the English proficiency test were able to perform above chance in the mismatch trials. Furthermore, figure 4.16 offers insight into the relationship between proficiency and the interaction between OI and phonological contrast. As already noted, accuracy was lower for the /f-v/ contrast words with both English and Arabic OI, compared to audio-only, whilst the opposite was true for the /m-n/ contrast words. Through the lens of proficiency, predictions for /f-v/ and /m-n/ contrast words were similar in the audio-only condition, even though higher proficiency with the /m-n/ contrast words would have been anticipated across all OI conditions. This implies that both phonological contrasts posed a certain amount of difficulty for the learners, which eased at a similar rate as proficiency increased. However, exposure to OI appears to have had a different effect on each contrast, as performance with the easier, well-established /m-n/ contrast was supported by OI, whereas performance with the more difficult, nonnative /f-v/ contrast was inhibited by OI. Notably, even those with the highest proficiency struggle to perform above chance when /f-v/ words are presented with any form of written input.

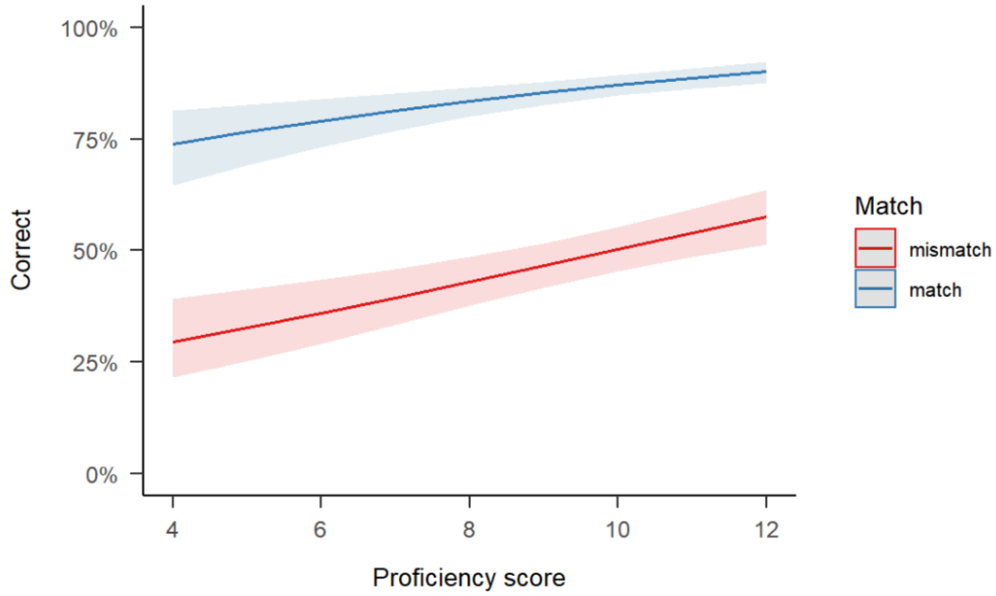


Figure 4.15: Model predicted probabilities for L1 Arabic accuracy, plotting trial type by L2 English proficiency

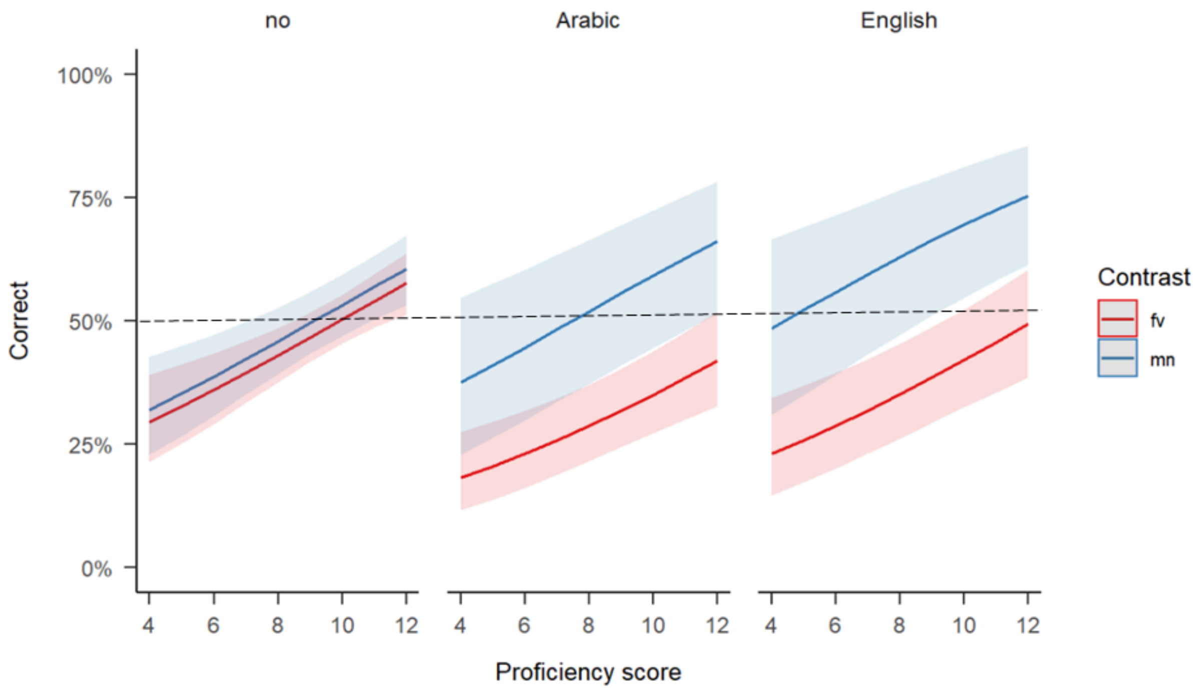


Figure 4.16: Model predicted probabilities for L1 Arabic accuracy plotting the interaction between phonological contrast and OI by L2 English proficiency

4.5.4 Word matching summary

The analysis of the L1 English-speakers' accuracy revealed expected main effects for trial type and orthographic input, where mismatch trials were more difficult and exposure to English spelling had a broadly beneficial effect. It was also unsurprising that accuracy reduced with age, especially considering the wider age range of the L1 English group and the demands of the task. The unexpected difficulty encountered with /f-v/ words prompted further investigation of the audio recordings and revealed partial devoicing of the voiced labiodental fricatives, which may have reduced perceptual salience of the voicing contrast. However, the primary concern was that L1 English-speakers were able to perform the matching task with above chance accuracy for the words differing by /f-v/, meaning that while the task was difficult it was possible. The interaction between phonological contrast and orthographic input, where accuracy was highest for /f-v/ words in the audio-only condition, further indicated that the contrast was sufficiently salient to lexically encode the contrast in memory, yet not sufficiently to benefit from the added orthographic input. Turning to the main analysis of the L1 Arabic group, the L1 English group's performance provided assurances that participants were able to learn the 12 bisyllabic pseudo-English words sufficiently to distinguish between minimal pair items in the subsequent audiovisual matching task.

The analysis of the L1 Arabic-speakers' accuracy revealed the predicted main effects for trial type, phonological contrast, and proficiency. Accuracy was lower with mismatch trials and words differing by /f-v/, meanwhile accuracy improved as L2 English proficiency increased. It was also unsurprising that the higher levels of reported distraction in the L1 Arabic group negatively affected task performance. Regarding the influence of OI, there was some indication that accuracy improved with English spelling compared to Arabic spelling. However, this was not well-supported by the data as a main effect and was not evident when looking at performance with /f-v/ words. The marginal advantage of L1 English written input across both contrasts may be indicative of Arabic-speakers' high levels of proficiency in English and a more macro-level congruence between script and English-like nature of the words. By congruence here, I do not refer to the mapping of sounds to letters between shared scripts, but rather where the presentation of L1 transliteration to L2 lexical items acquires a peculiarity and additional level of processing for biscriptal L2 learners who are used to learning new English words accompanied by English spelling.

For example, Hao & Yang (2021) had L1 English learners of L2 Mandarin with varying levels of proficiency complete a word learning and matching task, with real Mandarin words taught with either Romanised pinyin or Chinese characters. While target words were low-frequency and unlikely to be known to participants, the characters were familiar to both intermediate and advanced learners from their textbooks. Similar to the present study, they reported increased accuracy with higher proficiency and found that, at the advanced level, the character group generally outperformed the pinyin group. This tendency was also apparent for the intermediate learners, whereas naïve learners performed significantly better with pinyin than characters. The authors hypothesised that better performance with characters compared to pinyin for advanced learners was because this reflected learners' accustomed way of learning new vocabulary, as these literacy practices become more prevalent inside and outside the classroom as proficiency increases.

As mentioned, the interaction between contrast and orthographic input was not significant when comparing English and Arabic spellings, but only when comparing either of the OI conditions to the items presented with audio only. L1 Arabic-speaking participants distinguished words differing by both /f-v/ and /m-n/ at chance levels when they were taught without any accompanying spelling, which implied L1 Arabic-speakers struggled learning the words in general. It appears that those with the lowest levels of proficiency had not sufficiently learned the words to distinguish them in the matching task. However, across all learners the interaction between phonological contrast and OI vs no OI was apparent. Accuracy distinguishing words differing by the /f-v/ contrast was inhibited when taught with any OI, whereas accuracy with the easier /m-n/ contrast words was supported when taught with any OI.

These findings broadly align with those reported by Escudero (2015), who reported that OI only positively influenced L2 word learning and recognition when novel word pairs differed by a contrast that could already be perceived, thus functioning as a confirmatory cue. No effect of orthography was found for the most difficult to perceive contrasts or, in fact, the easiest to perceive contrasts either, where OI was argued to be redundant. She also found that this pattern was evidenced across proficiency levels, as in the present study. It was somewhat surprising that the /m-n/ words did not mirror the findings for perceptually easy items in the Escudero (2015) study. However, one of the reasons why /m-n/ items may not have been as easy may be the added difficulty of learning bisyllabic items in the present study rather than monosyllabic items in Escudero's study. Interestingly, when looking at the same interaction with the L1 English group, the influence of OI with /m-n/ items is far less apparent, more closely resembling the proposals of Escudero (2015). Additionally, the added difficulty observed for the L1 English group with OI when accompanying /f-v/ items may extend Escudero's findings to contrasts that are broadly difficult to acoustically distinguish, rather than specifically nonnative phonemic contrasts. However, this line of enquiry is beyond the scope of the present study.

Overall, the findings reported above do not provide evidence that either shared or distinct written input supported the learning or recognition of words differing by a difficult to perceive phonological contrast. Additionally, while there was some indication that L1 Arabic learners of L2 English broadly benefited from exposure to English spellings, more than Arabic spellings, this was not well-supported by the data and did not lead to improved performance with /f-v/ contrast words. The concluding section reviews the originally stated research question and predictions at the start of the chapter in light of these results.

4.6 Concluding remarks

This chapter addressed the question: to what extent do different types of orthographic input (OI) influence the lexical encoding of L2 phonological contrasts in memory during novel word learning? L1 Arabic-speakers' ability to lexically encode the distinction between pseudowords differing by a confusable L2 contrast (/f-v/) was assessed in an audio-visual matching task, after a short word learning session. Based on the relevant literature, predictions were made regarding the perceptibility of the target contrast, congruence of shared OI, systematicity of distinct OI, and L2 English proficiency. When looking at the performance of the L1 Arabic group with words that differed by /f-v/, there was insufficient evidence to suggest that English OI had a different influence on the successful encoding of the target contrast compared to Arabic OI. However, that is not to say there was no influence of OI, as written input in both scripts was found to have a inhibitory effect on lexical encoding, when compared to words presented without any OI.

Looking at the specific predictions in more depth, most anticipated effects were evidenced in the reported findings. L1 Arabic-speakers responded with lower accuracy and slower response times with words differing by the confusable /f-v/ contrast, compared to the more perceptually salient /m-n/ contrast. Surprisingly, this extended to the L1 English group, highlighting the need to consider the ecological validity of learning new minimal pair items, accompanied by novel images with low nameability, and differing by a nonnative phonological contrast which exhibited low perceptual salience even for L1 English speakers. To what extent does the difficulty posed by this task reflect experience of language learning outside of an experimental setting? Participant performance with the easier contrast showed that the task itself was achievable and participants in both groups were able to perform with a high level of accuracy. Additionally, several participants provided informal feedback that they continued thinking about the images and lexical items long after completing the study, as well as suggesting they found both the words and images to be fun. Chapters 5 and 6 both offer additional insight into participant reflections and their experiences of learning the words in study. Returning to the present question, performance and feedback implied that participants were able to learn and recognise the words within a short experimental session. However, the perceptual salience of target contrasts was a significant factor in relation to overall performance, and specifically the influence of orthographic input.

It was predicted that exposure to Arabic OI would lead to interference effects for /f-v/ words, due to the incongruent-shared written representation of the two sounds with a single letter in Arabic. Lower accuracy and slower response times were recorded when words had been accompanied by Arabic spelling. However, performance did not differ significantly from words accompanied by English spelling. This was in opposition to the hypothesis that the systematic-distinct English OI would have a facilitative effect on distinguishing /f-v/ words, through the visual analysis of the confusable contrast. Instead, it appeared that both forms of written input had an inhibitory effect. These findings were aligned with proposals by Escudero and colleagues (2015; 2014) that argued the supportive influence of OI was limited by the perceptibility of target contrasts, and functions to enhance what learners can already perceive. They reported a null influence of orthography, rather than an inhibitory effect of OI that was familiar, congruent, and relatively easy to perceive. However, an inhibitory effect was reported by Mathieu (2016), who compared participant performance across a range of different script conditions. He found that written input conditions, which varied in familiarity based on the L1 script, did not differ from each other but together exhibited an inhibitory effect, compared to words taught without any OI. This was attributed to the "foreign-ness" of the written input, as none were fully in the same script as the L1. While the English OI could be conceived of as "foreign" to the L1 Arabic-speakers, these ill-fitting terms for the present study illustrate the benefit of articulating the degree of overlap between scripts on the spectrum of shared-distinct, rather than familiar-unfamiliar.

The final prediction was that performance would improve as L2 English proficiency increased. This was indeed evidenced as a main effect, and was robust across phonological contrast and OI conditions. While participant accuracy distinguishing the /f-v/ words gradually improved with proficiency, only those with the highest scores on the proficiency test performed above chance. In terms of the interaction between orthography and phonology, it was found that, at all proficiency levels, participants performed worse when /f-v/ words had been taught with any OI. There was evidence that as proficiency increased participants were better able to navigate interference from Arabic incongruent-shared written forms and integrate the English systematic-distinct written forms in memory. However, even participants with the highest proficiency scores struggled to perform above chance when /f-v/ words were accompanied by any OI.

The fact that all participants had some familiarity with English spelling, and that participants with high English proficiency were also highly literate in English, indicates that the effects reported here were not related to familiarity of OI. Instead, these findings reflect the increased difficulty of lexically encoding a confusable contrast with sufficient phonetic detail when any OI is available. As the opposite was true with the easier /m-n/ items, it appears that orthographic influence depends on the perceptual salience of target items. Thus, it is hypothesised that when learners are able to perceptually distinguish lexical items, OI functions as an enhancing cue that supports lexical encoding. However, when learners are not able to perceptually distinguish novel lexical items, OI causes interference or distracts from attending to the necessary phonetic detail. In the next chapter, analysis of participants' reflections, related to orthographic influence when learning words across writing systems, sheds further light on these findings.

Chapter 5: Participant perspectives on written input

5.1 Introduction

This chapter presents the analysis of post-test questionnaire responses, designed to explore how exposure to different types of written input were perceived to influence novel word learning. Questions in the first section of the post-test questionnaire combined open and multiple-choice responses to investigate perceptions of the influence of orthographic input (OI) during the word learning and matching tasks, as well as broader beliefs about the importance of OI when learning new words. The data from the two language groups are analysed separately and then discussed comparatively within the categories that are formed inductively from the data. As before, the L1 Arabic-speakers are the focus of the analysis outlined below. The research questions being addressed are:

What influence of written input were learners aware of during word learning?

1) *What was the perceived influence of written input during the present study?*
2) *What were the participants' beliefs about the importance of written input when learning new words more broadly?*

To date, there is little research within the field of orthographic influence on L2 phonological acquisition that includes participant perspectives, or any form of qualitative analysis. Indeed, the literature reviewed thus far has almost exclusively focused on behavioural evidence for the influence of orthographic input on L2 learning and processing. However, one study by Bassetti and colleagues (2020) has included qualitative analysis of participant perceptions, when investigating the influence of L2 orthographic forms on speech production and phonological awareness of L1 Italian-speaking learners of L2 English. A questionnaire featured in their battery of tests, which included an open question where participants were invited to explain their responses on the phonological awareness task. Although not explicitly stated, the qualitative analysis of questionnaire responses seems comparable to Qualitative Content Analysis (QCA), where categories were inductively developed, deductively applied and combined with descriptive quantitative frequencies.

> Responses were transcribed and coded using one of the five categories that emerged from the qualitative analysis. Categories that described the target sound were "double", "long", "strong/harsh", or "target sound (other)". Responses that described sound differences other than the target sound were categorized as "other" (e.g. word stress). (Bassetti et al., 2020, p. 1236)

These results were then combined with quantitative findings to illuminate why participants incorrectly rejected L2 English rhymes, in relation to their interpretation of the orthographic input. Qualitative insights also shed light on how participants conceptualised the difference between phonological contrasts. For example, 92% of those who incorrectly rejected the rhyme *very-merry*, and provided a valid explanation, said that the /r/ in *merry* was "double", "long", or "stronger/harsher", demonstrating how qualitative data can complement and further explain quantitative findings. Overall, the qualitative data in the study confirmed that the main reason that participants rejected rhymes in English was consonant length, and that this was the same contrast that participants applied to Italian words, such as *caro-carro*.

In a different line of research, related to collaborative L2 writing with technology, Steinberger (2017) offers a more explicit example of QCA with data very similar to that found in the present study. While his sample was smaller, with only 24 participants, his doctoral thesis shows how QCA can be applied to short survey responses and used to inductively form categories for exploratory studies, following procedures outlined by Mayring (2014, 2015, 2019). Steinberger presents a transparent overview of how he created inductive categories and coded his data, referring to this process as 'tagging'. He went on to count, cluster, and visualise his tags, which provided a useful overview of both general and specific content from the data, as well as quantitative and qualitative insights. For example, participants' responses to the survey question "How did the synchronous nature of the task affect your group work?" were first divided into two categories *positive* and *negative*, then further subdivided into more specific categories. These included six positive categories: *general*, *efficiency*, *multimodality*, *transparency*, *language-related peer feedback*, and *idea generation*; and three negative categories: *no face-to-face*, *confusing*, and *time-consuming*. This revealed that participants generally perceived synchronous computer-mediated collaboration to be positive (84.4% of tags) and even though most positive tags were *general*, there were a range of more specific ways that participants perceived this impact, most notably *efficiency*.

The present study similarly seeks to gain insight into participants' perceptions of their learning and processes through short responses to open questions and qualitative analysis, demonstrating the advantage of mixed methods to triangulate behavioural findings. The sections below review the method used and outline the analytical approach, before presenting the results related to perception of OI influence on the task, as well as beliefs around OI importance more generally. Thus, the present chapter aims to better understand how participants believe their learning is impacted by written input. Before moving to the methods and analysis, it is worth returning to my ontological and epistemological position, which is outlined in more detail in chapter 3.1.

Regarding the analysis of both quantitative and qualitative data in this study, I assume an active role in *making* sense of patterns based on partial evidence. Whether it be through constructing statistical models or forming codes and categories, I assume that the analytical steps I take and conclusions I arrive at are influenced by my theoretical background and personal axiology. In both cases, I strive to adopt an approach that is rigorous, reproducible, and reflexive. My position remains post-positivist and my approach to qualitative research is likely to be characterised as "small q" (Braun & Clarke, 2013), due to the predominance of broad numeric data, focus on general patterns and interest in explanatory or predictive relationships between variables in this thesis. However, I believe this to be coherent with approaches such as those advocated by Mayring (2014, 2015, 2019), who outlines ways to integrate quantitative insights into QCA and make systematic use of both deductive and inductive approaches.

In the two studies exemplifying possible applications of a QCA approach, both Bassetti (2020) and Steinberger (2017) refer to categories "emerging". However, many qualitative researchers take issue with this phrasing, most famously Braun and Clarke (Burnage, 2021), as it can imply that themes and categories are ontologically real and materially exist, waiting to be passively discovered. In agreement with this criticism and in line with Mayring's own wording that a "a category has to be constructed" (2014, p. 81), I use language that suggests actively and iteratively creating, shaping and honing categories to interpret the data and offer a truthful account. I acknowledge the influences that I bring to the interpretative process, which inevitably shape the lens through which I understand participant responses. Predominantly, these influences include my perspective as a

Western-educated, English-speaking, British national and my experiences as a language researcher, teacher and learner (see 3.1.1). I mention this with the aim of transparency and reflexive rigour rather than emphasising subjectivity, as reproducibility and agreement across multiple coders is also valued as a measure of valid and reliable interpretation.

5.2 Methodology

For clarity, the methods applied to address the research questions stated at the start of this chapter are reiterated here. Immediately after completing the word learning and matching tasks, participants were presented with a post-test questionnaire, which invited reflections on the influence of written input and how participants had consciously attempted to learn the words. The questions relevant to the analysis in this chapter are found in table 5.1.

Table 5.1: Overview of post-test questionnaire items relevant to learner perceptions of OI influence and importance when learning new words

Focus area	Question	Type
Perception of OI influence during the study	1) When did you first notice the English or Arabic spelling?	Multiple choice *[beginning, middle, end, I didn't see any spellings]*
	2) Seeing the words written in Arabic letters…	Multiple choice *[made it easier, made no difference, made it more difficult, I didn't notice]*
	3) Seeing the words written in English letters…	Multiple choice *[made it easier, made no difference, made it more difficult, I didn't notice]*
	4) Hearing the words without the spelling…	Multiple choice *[made it easier, made no difference, made it more difficult, I didn't notice]*
	5) I would have preferred to hear the words and see all of them…	Multiple choice *[written in Arabic, written in English, without any spelling]*

	6) Did it make a difference seeing the written words? In what way?	Open
Importance of early exposure to OI in word learning, in general	7) In general, when learning new words do you think it is important to see the spelling the first time you hear it?	Binary choice *[yes, no]*
	8) Why?	Open

The responses to questions (Q) 1-6 are described and discussed in section 5.4, while Q7 and Q8 are analysed in section 5.5. The questions have been divided in this way to relate to the two research sub-questions, which are the topic of this chapter. For example, section 5.4 addresses perceptions of orthographic input, specifically in relation to the experimental task. Meanwhile, section 5.5 addresses beliefs about the importance of early exposure to written input during language learning more generally.

The first questionnaire item was added after piloting revealed some participants did not notice the written input below the images, or started to notice the written forms midway through the presentation trials. This is important to ascertain, as the subsequent questions assume that participants were consciously aware of the written input they were exposed to. Q2-4 compare the perceived difficulty of learning new words with different types of written input. Perceived difficulty is a common measurement of cognitive appraisal in SLA research, mostly to confirm methodological operationalisation of task complexity, but also to explore emotional responses such as stress and interest (Cho, 2018). Participants were then asked in Q5 to explicitly state their (written) input preference, as it is assumed that something can be perceived to be difficult but also preferable or helpful. The results of these quantitative measures are presented briefly and descriptively in section 5.4.1.

In order to deepen understanding of these findings, Q6 and Q8 were included to elicit open responses from participants. Responses to the open questions were typically short, constituting one to two sentences. As with all the study content, participants were provided with the questions in their first language; however, they were able to respond to open questions in the language of their choosing. This resulted in a mix of English and Arabic responses from the L1 Arabic group. Meanwhile, the L1 English group only responded in English. A sample of the responses from both groups for each open question are provided in table 5.2 and 5.3, to offer some insight into the type of data elicited.

Table 5.2: Random sample of responses to Q6 by both L1 Arabic and English-speakers

Q 6: Did it make a difference seeing the written words? In what way? **عندما تتعلم كلمات جديدة هل أحدث ذلك فرقا في رؤية الكلمات المكتوبة؟ إلي أي مدي؟**		
Group	**Participant ID**	**Response**
Arabic	5442543	بعض الكلمات لم ألاحظ أنها تبدأ بصوتين متشابهين مثلا v and f وبما أن هذه الأصوات غائبة في لغتي في الغالب لن أنتبه لها عند سماعها فقط فالإملاء يساعد في هذه الحالة [*I didn't notice that some of the words started with two similar sounds, for example v and f. Since these sounds are mostly absent in my language, I won't pay attention to them when I hear them, only the spelling helps in this case.*][1]
Arabic	5015277	yes ,i prefer the word written while listening to it, that's how i memorize the word twice written and spoken
Arabic	5357676	نعم فقد ساعدني في تذكر شكل الكلمة كثيرا سواء الكلمة بالعربية او الانجليزية فأنا أتخيلها واتذكر لو حتى حرفا واحدا [*Yes, it helped me a lot with remembering the shape of the word, whether the word was in Arabic or in English, so I imagine it and remember, even if only one letter.*]
Arabic	5336601	writing words in Arabic made it somehow challenging to learn them.
Arabic	5338948	جعل تذكرها اسهل [*It made it easier to remember it*]
English	5304273	yes - easier to help remember when in english
English	5336201	it helps to picture the word when looking at the picture so they become like a pair in my mind
English	5308502	i think it aided the distinction between similar sounding words
English	5312160	It makes a difference to me in general because I can see the sounds I am unable to pick up by just listening.
English	5521508	I preferred learning from sound alone, as English spelling didn't reflect pronunciation.

[1] Translation offered in brackets and italics for the Arabic language responses

Table 5.3: Random sample of responses to Q7-8 by both L1 English and Arabic-speakers

Q7: In general, when learning new words do you think it is important to see the spelling the first time you hear it?

بشكل عام، عندما تتعلم كلمات جديدة هل من المهم عندك أن ترى املاء الكلمة عند الاستماع لها لأول مرة ؟

Q8: Why?

لماذا؟

Group	Participant ID	Response	
Arabic	5499647	no	لكي لا انشتت بالتركيز على املاء تلك الكلمة [*In order to not get distracted by focusing on the spelling of that word.*]
Arabic	5059373	yes	i am a visual learner and seeing the words written help me to memorize better
Arabic	5196532	yes	mental connection between letters and how the sound. make sure about the right sound. for example, bin or ban. to make sure i got the right word
Arabic	5339892	yes	حتى أفرق بينها وبين كلمة أخرى شبيهة ، وحتى أتعرف عليها في سياق القراءة، ولأتمكن من استخدامها في التحدث والكتابة. [*So I differentiate between it and another similar word, and so I recognise it in the context of reading, and I can use it in speaking and writing.*]
Arabic	5201938	yes	to remember the word and know how to write it
English	5304273	no	easier to hear and then understand how its written
English	5499421	yes	I think I find it quite difficult to make the correct sounds when speaking, without first seeing the word written down (assuming it is written in characters I am familiar with)
English	5338550	yes	to confirm youve heard the sounds correctly
English	5332693	yes	Visual learner, my memory seems quite dependent on having something in front of me to help with learning
English	5659326	no	In english the spelling doesn't always match up with pronunciation so it could make it difficult.

5.3 Analytical Procedure

As the aim of this chapter is to gain a better understanding of participants' perceptions of how written input influences their learning, the emphasis of the analysis and discussion is on the qualitative content analysis (QCA) of the open responses to questions 6 and 8. This analytical approach was chosen for its philosophical flexibility and attention to both description and interpretation of data, considering context and seeking themes (Vaismoradi & Snelgrove, 2019). There are many similarities between thematic analysis and QCA, where arguably either could appropriately address the question and aims at hand (Braun & Clarke, 2021; Neuendorf, 2018; Vaismoradi & Snelgrove, 2019). However, the roots of QCA in a more post-positivist paradigm, combined with the research question focusing on the "what" and manifest content of participant responses, led to the decision to follow this framework. QCA also offers the opportunity to include frequencies and quantitative descriptions as intermediary steps, and a form of mixed methods data triangulation (Mayring, 2014). Thus, brief descriptions of frequencies are provided for the multiple-choice questions at the start of section 5.4 and 5.5, as well as category frequencies to provide additional context for the qualitative analysis.

The QCA approach here follows a similar procedure to that laid out by Mayring (2014) for inductive category formation, summarised in figure 5.1. This approach to QCA was chosen for its focus on material relevant to the research question(s) and arrival directly at summarising categories that are tightly connected to the material itself. Mayring (2014) notes the similarity between this process and "open coding" within Grounded Theory (A. Strauss & Corbin, 1990; A. L. Strauss, 1987), but argues that the QCA approach needs to be more systematic and this can be achieved by following the steps stated in figure 5.1.

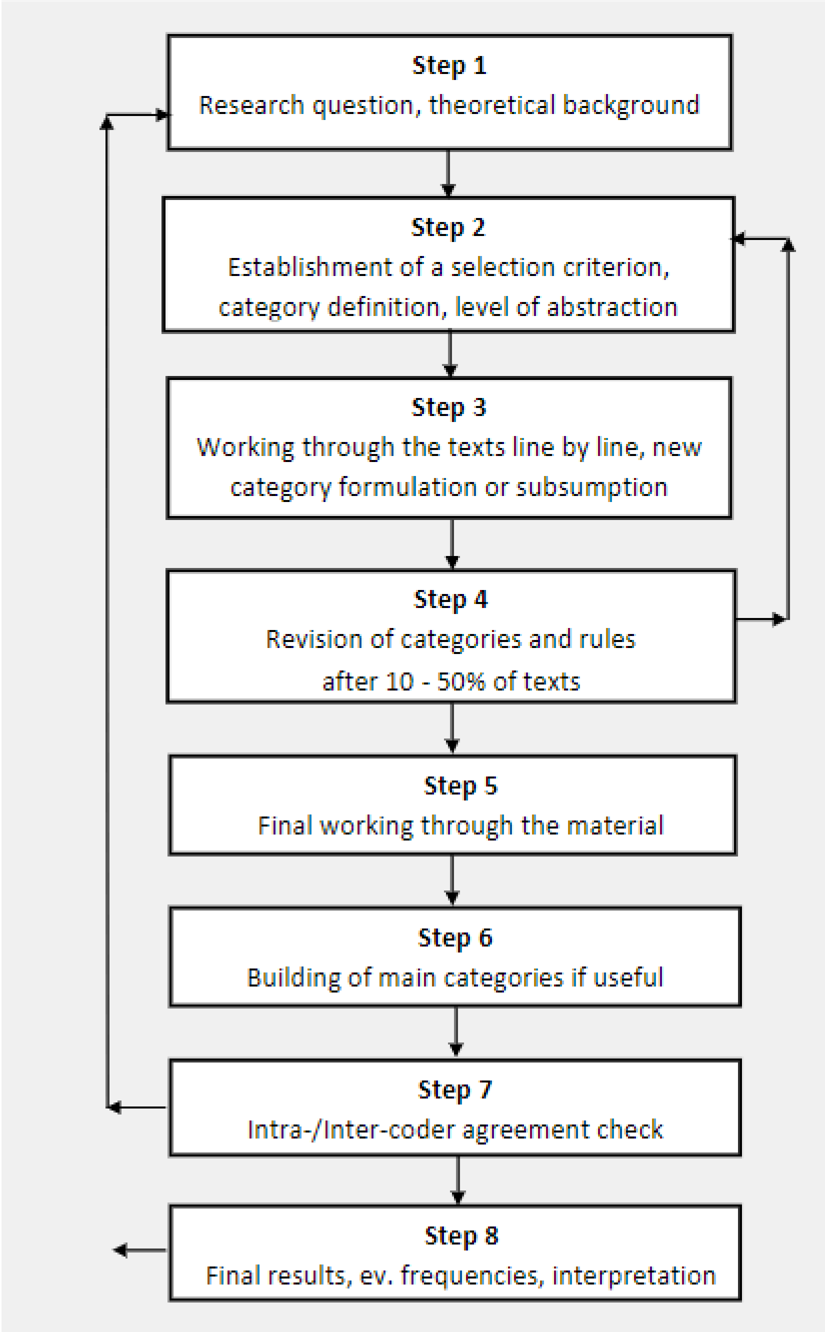


Figure 5.1: Steps of inductive category development (Mayring, 2014, p. 80)

The questionnaire material relevant to this section of analysis was oriented towards understanding participants' awareness of orthographic influence on the learning of new words, and encouraged participants to express how they experienced the different types of written input that were presented to them. Therefore, the text is used to arrive at statements on participants' perceptions of types of influences related to written input during word learning in the study. In accordance with step 2, existing research and relevant theoretical considerations were used to formulate selection criteria, category definitions and level of abstraction before analysis. Table 5.4 shows how these were defined for the material in relation to both Q6 and Q8. For example, for Q6, academic debate around the topic of OI influence on L2 phonology has often focused on understanding whether written input has a broadly positive, negative or neutral influence on L2 phonological development There is also evidence that familiarity with the writing system of the written input is influential, as well as the systematicity of different orthographies (Hayes-Harb & Barrios, 2021). Meanwhile, Q8 draws on research into learner beliefs, which has shed light on not only *what* beliefs learners hold but also *how* these beliefs are formed, change, and interact with the learning process (Barcelos & Kalaja, 2011).

Table 5.4: Overview of selection criteria, category definitions and level of abstraction for analysing responses to Q6 and Q8

	Q6	**Q8**
Coding unit	Clear semantic units within participant open responses.	
Context unit	Whole response to relevant question, the wording of the question, responses in the background and debrief questionnaire broadly.	
Recording unit	Questionnaire responses for 117 L1 English-speaking[29] and 114 L1 Arabic-speaking participants	
Category definition	● Positive, negative or neutral perceptions of orthographic influence. ● Perceptions differing between orthographies. ● Perceptions of influence connected to language processing and learning.	● Affirmative or negative views towards the importance of OI when learning new words. ● Specific focus on first exposure and related to different script OI.
Level of abstraction	● General evaluation of OI influence and whether there is any difference between scripts. ● Specific insights into types of influence on processing or learning.	● General beliefs regarding early OI importance and impact on early language learning and processing. ● Specific insights into beliefs and justifying OI (lack of) importance.

[29] The questionnaire open responses revealed that one L1 English participant was able to read in Farsi, so they were also excluded from the remaining analysis.

As mentioned, all the L1 English-speakers responded in English, whereas around half of the L1 Arabic group responded in Arabic and the other half responded in English. As the language choice was not always consistent across the open questions, it appears participants were more comfortable expressing themselves on different topics in different languages. Responses remained in their original typed form and were not adjusted or translated before coding. This decision was based on the desire to maintain transparency in the analysis, stay close to the original content, and to keep the nuance of the original responses as intact as possible. As will be further explained, the codes themselves were all in English, for continuity across the whole dataset and drawing on my stronger linguistic resources in that language.

The next step in the process was to read the material line by line, in relation to the category definition, and to decide whether it should be included or ignored. Then, a category was formulated at the outlined level of abstraction, with efforts made to stay close to the original wording of the text. The next passage was then assessed to see whether it fitted within the first category or required the formulation of a new category, and so on for the next passages. First, I coded the passages from the L1 English-speakers' pilot data (n = 5) for Q6, to assess whether the category definition and level of abstraction were appropriate, and made necessary revisions. I then coded the first ~50% of responses which resulted in around 12 positive influence categories and three negative influence categories. These were revised down to 10 positive influence categories and two negative influence categories, which were then applied to the rest of the responses following the same procedure as before.

For the second run through, each response was printed, cut out, and colour-coded for whether the overall evaluation of OI during the study was positive, negative, or neutral, as well as specific reference to English or Arabic script input. Then, responses were coded according to the categories devised during the first run-through, in relation to any specific influence of OI on language processing or learning. Next, responses were grouped into their categories and closely compared to assess the validity of the categories and whether they need to be revised. At this stage, main categories and subcategories were established where useful. After completing the second run through of the English questionnaire responses, I turned to the Arabic questionnaire responses to see whether the codes would extend to the data from this language group. While there may be different perspectives as to whether it is appropriate to derive categories based on the English data and then apply the same codes more deductively to the Arabic data, this appeared the best route for three pragmatic reasons. Firstly, as English is my first language, I was better able to engage with the nuance of the English responses initially and form suitable English category labels. Secondly, a larger proportion of the L1 English group gave longer and more detailed responses, thus the data offered a somewhat richer starting point. Thirdly, around 50% of L1 Arabic-speaker answers to both Q6 and Q8 were written in English, so for these responses, at least, it was logical to extend the application of English language categories.

In order to avoid forcing the Arabic responses into potentially inappropriate categories, I took the following approach to collaboratively code the L1 Arabic responses with an L1 Arabic-speaking researcher in my department. As with the English responses, I coded all the responses continuing with the process of QCA, whereby each response was assessed in terms of each previously formed category and where necessary new categories were created. After the first pass, I cut up all the responses and grouped them together with the other responses that I believed fell within the same category. I then went through each group of responses with my colleague and asked her 1) whether she agreed that these responses all belonged to the same category, and 2) what she would call that category. We then compared her categories to mine and discussed the best phrasing together, as well as how to address the new categories that were not represented in the L1 English responses.

There was a high level of agreement between us, as she agreed that all the responses I had grouped together belonged to the same category and table 5.5 demonstrates the commonality between our category labels. Through comparison and discussion, it was decided that I should continue with the categories that I formulated and it was agreed that an additional broad category for *helps learning* should be added based on the Arabic data. It was also noted that none of the L1 Arabic responses were a good fit for the categories of *focus on the first letter*, *rehearse the word* or *unfamiliar*, which are further discussed in the results. Taking the L1 English and L1 Arabic data together, there were five positive influence categories, as well as five subcategories, and three categories where OI was not perceived to be helpful, including both negative and neutral responses. These are described and discussed more fully in section 5.4.

Table 5.5: Examples of category agreement between coders for Q6

My categories	L1 Arabic collaborator categories
Learning	Makes learning easier
Remembering	Remember
Connect audio and visual input	Connect audio and visual
Map sounds to letters	Connect written and spoken form
Associate word and image	Associate image and written form
Visualise words	Visualise word
Clarify what I heard	Hear correct form
Differentiate similar sounds/words	Difference between written and spoken form (similar sounding words)
Distracting (from sounds)	Visual distraction
Better to focus on sounds	Focus on hearing

The same procedure was followed with responses to Q8, with the only difference being that no general evaluation needed to be inferred, as participants had provided yes-no responses for broad categorisation of whether they felt OI was important or not when learning new words. Therefore, the focus was on the specific categories for justification of this belief, including ways that seeing written forms impacts the memory, processing, and general learning of words in a new language. Coding the first ~50% of L1 English responses resulted in around 14 categories related to why OI was seen to be important when learning new words, and then eight categories related to why early OI exposure was not seen to be important. After the second run through, these categories were revised to eight categories for early OI importance and four categories for lack of importance. These categories were then applied to the L1 Arabic responses, following the same collaborative procedure as outlined above. The categories based on the L1 English responses were broadly a good fit for the L1 Arabic data. However, two new categories needed to be formed and it also became clear that some of the existing categories should be revised and reduced. This resulted in seven categories for early OI importance, including one category with three subcategories, and four revised categories for the lack of early OI importance, which are detailed and discussed in section 5.5.

Coding guidelines were given to two additional coders for each language group, alongside a sample of 20% of the questionnaire responses. Two L1 English colleagues were given data from the English questionnaires and two L1 Arabic colleagues were given data from the Arabic questionnaires.[30] All coders had a background in Linguistics and were either doctoral candidates or had completed their PhDs at the University of York. Coders were instructed to assign the formulated categories to the sample of responses, but were also able to leave additional comments or ask questions. This additional dialogue between myself and the coders ensured understanding of the approach and the categories to be assigned. The materials and instructions shared with coders via email can be found in appendix XI and XII. Intercoder agreement was calculated using Krippendorff's Alpha (α), which is a statistical measure of agreement, or more accurately disagreement, that is applicable to multiple coders, categories, or values, rather than being limited to measuring the agreement between two coders/judges (Hayes & Krippendorff, 2007; Krippendorff, 2017). It can also handle different levels of measurement and missing data well, which is part of the reason for the popularity of this calculation for QCA intercoder reliability measurement (Gläser-Zikuda et al., 2020; Mayring, 2014; O'Connor & Joffe, 2020). The R script used to calculate the Krippendorf's Alpha for Q6 and Q8 coding can be found in appendix XIII.

5.4 Participant perceptions of OI influence during the task

As there is very little literature regarding participants' perceptions of learning new words with different forms of input, especially across writing systems, this analysis is exploratory and there are no explicit hypotheses. In the previous chapter, participants' accuracy scores on a matching task revealed both L1 English and Arabic groups exhibited worse performance with words differing by /f-v/ compared to /m-n/, where Arabic speakers particularly struggled to perform above chance. For both groups, there was an overall advantage for correctly identifying items that had been taught accompanied by English

[30] One L1 Arabic colleague misunderstood the task, which led to her mostly generating her own categories. These did overlap in most cases, but to maintain consistency in approach and number of coders, my previous collaborator was invited back to re-code the 20% sample. This meant that we had discussed the codes and categories together in more depth than was the case with the other coders.

spelling. However, this did not significantly contribute to the model for the L1 Arabic group and did not extend to words differing by confusable contrasts. In fact, there was an inhibitory effect of any OI, either English or Arabic, when words differed by the difficult /f-v/ contrast. In addition, it was found that, in the L1 Arabic group, L2 English proficiency was a strong predictor of task performance. The reasons why both English and Arabic OI had an inhibitory effect, even with high levels of proficiency, remains unclear. Furthermore, the suggestion that the different script inputs do not have distinct influences on the learning of confusable contrasts is surprising and deserves more attention. The exploration of participant self-reported perceptions in this chapter offers further explanation of participant processing and task performance in relation to OI exposure.

Based on the results from chapter 4, it is anticipated that participants from both L1 groups will indicate a preference for learning with English spelling. It is also expected that participants in both groups would note that /f-v/ words were particularly difficult. It is then of interest whether participants note inhibitory influences of OI, specifically in relation to this phonological contrast. As in the previous chapter, the L1 English group findings are not reported for comparison to a native-like target. Instead, the L1 English questionnaire responses are explored to shed light on the perceived influence of OI when phonological contrasts are found in the L1, but represented by unfamiliar and familiar written forms. Meanwhile, the L1 Arabic questionnaire responses relate to the perceived influence of OI when learning a difficult nonnative phonological contrast, in the context of biliteracy.

5.4.1 Frequencies of perceived OI influence

It was of interest to know whether people noticed the written forms and at what point in the learning phase, as some people mentioned during piloting that they did not see the written words at all. Figure 5.2 indicates that most participants noticed the words at the start or at least the middle of the word learning trials[31]. However, while a minority of participants of L1 English-speakers reported not seeing the written words at all, a much larger proportion of the L1 Arabic-speakers selected this option. This amount is peculiar and may indicate ambiguity in the Arabic translation of the question. However, based on open responses, at

[31] As trial presentation was randomised and a third of the trials presented words without OI, it is possible that some participants were not presented with written forms until midway through.

least some participants believed they had not seen any written forms, evidenced in the following quotes:

"I don't remember seeing any written words! I feel a bit silly now." (EN5588274[32])

"I only saw picture and heard words…well there would be no point in seeing arabic for me. i didnt see any words at all."(EN5663873)

The second quote was excluded from the analysis, as the speaker was over 65, indicating possible age-related limitations. However, the first quote came from a 31-year-old, which undermines that logic. As this was also reported in piloting, when I was present and could confirm that the written forms had appeared, this was unlikely to reflect a technical error. Therefore, it is worth considering the possibility that some participants were not consciously aware of having seen any written input. That being said, several open responses of those who reported "not seeing" the written input do mention seeing the different spellings, implying a misunderstanding of the question.

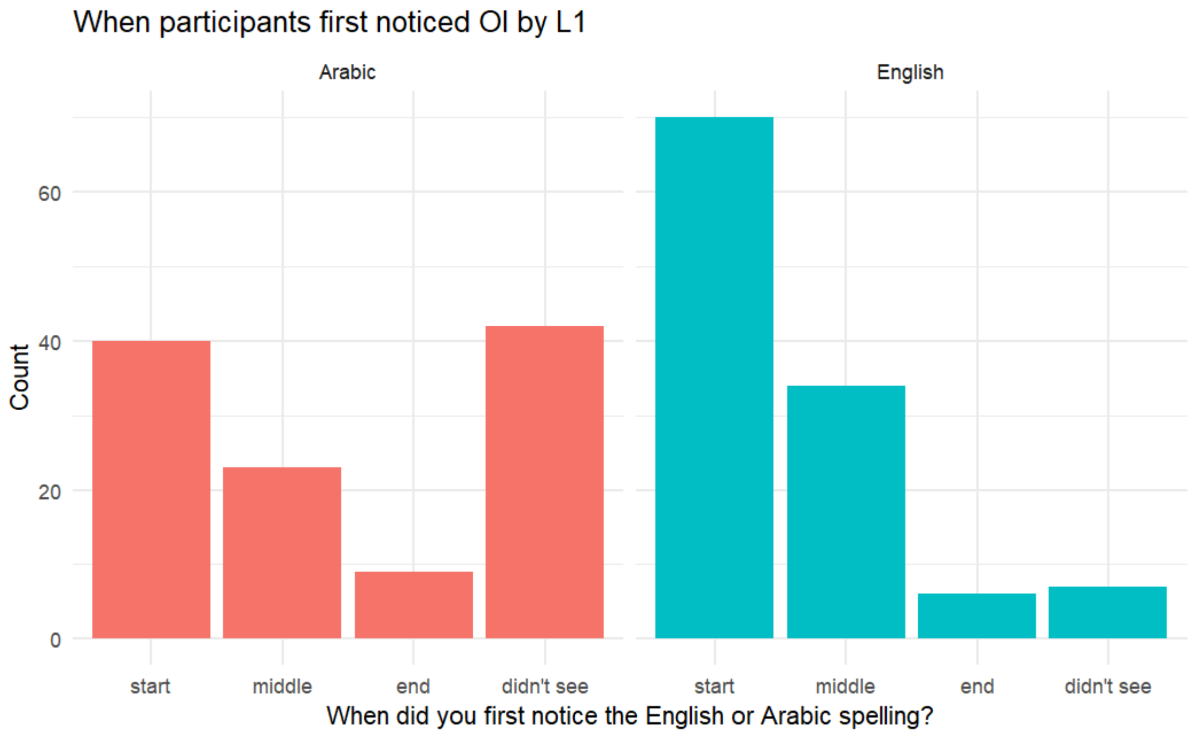


Figure 5.2: Questionnaire responses for when OI was first noticed during the learning phase of the experiment, by L1 group

[32] Participant ID numbers accompany the quotes and are given either an EN or AR prefix depending on whether the participant was an L1 English or Arabic-speaker. The same numbers are used to identify participants throughout all data files and appendices, for transparency.

When asked about the perceived difficulty of learning with different types of input and overall input preference, figure 5.3 demonstrates that both groups reported a strong preference for seeing the English spelling. The second most popular option for the L1 English group was learning without any written input, whereas the L1 Arabic group preferred learning with the Arabic spelling. Echoing these preferences (figure 5.4), most participants from both groups reported that English OI made learning easier, with a small proportion of the L1 English group reporting that it made no difference. Then, as would be predicted, the L1 English group reported that exposure to unfamiliar Arabic OI mostly made no difference, but in some cases made it more difficult. Responses were more varied in the L1 Arabic group, where a slight majority found learning easier with shared Arabic OI, while others found it more difficult or said it made no difference. Finally, similar proportions of the L1 English group reported that the audio-only presentation made it more difficult for them to learn or made no difference, with only a small number reporting that it made learning easier. In contrast, a clearer majority of the L1 Arabic group reported that audio-only input made learning more difficult.

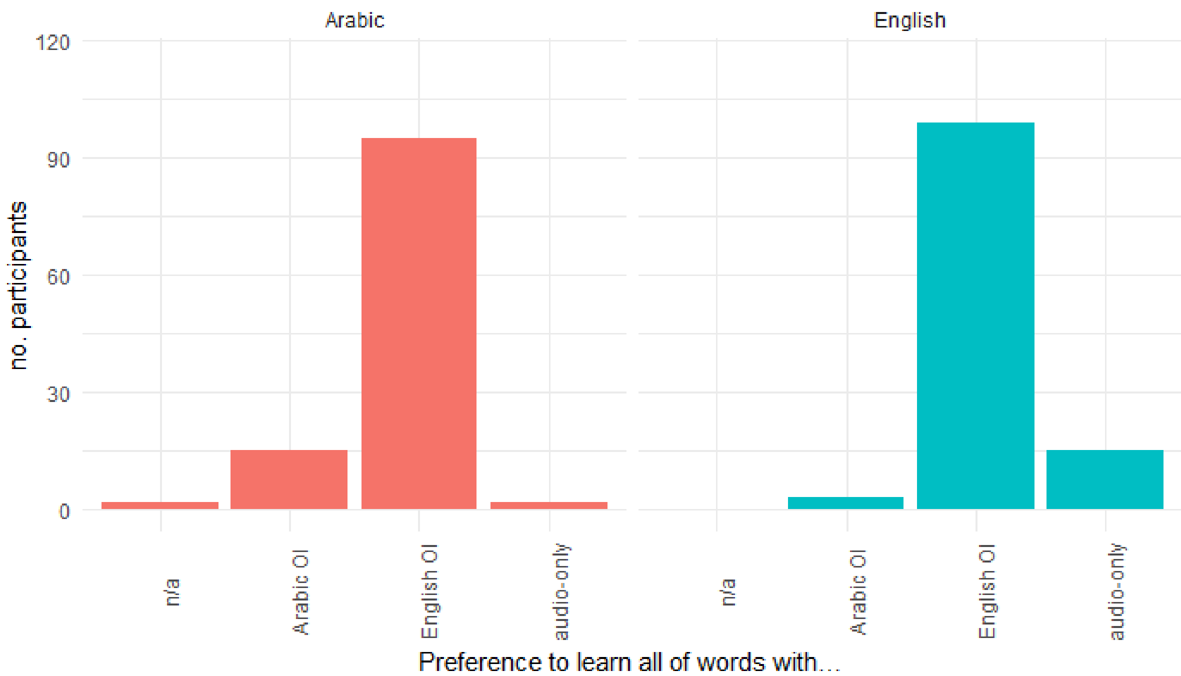


Figure 5.3: Questionnaire responses for overall spelling preference, by L1 group

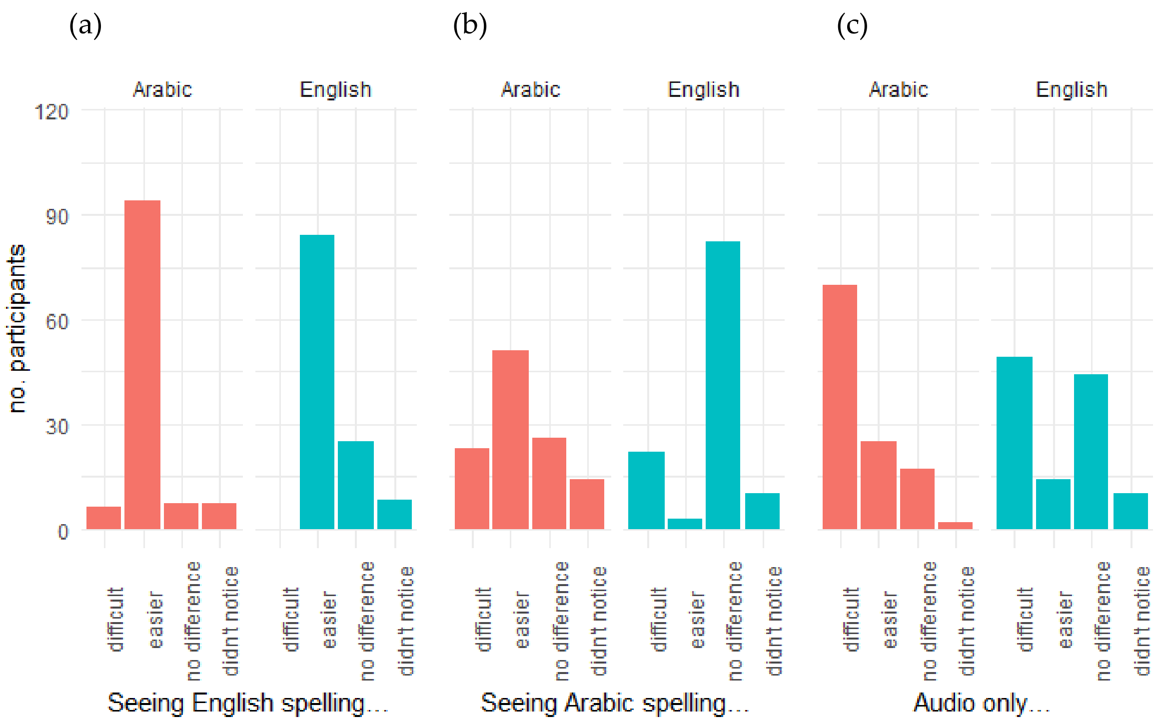


Figure 5.4: Questionnaire responses for perceived influence of (a) English OI, (b) Arabic OI, and (c) no OI, by L1 group

These responses were briefly explored in relation to proficiency test scores for the L1 Arabic-speakers (figure 5.5 and 5.6). Overall, the preferences are quite clear, irrespective of proficiency level. However, it is noteworthy that those who reported that English spelling was more difficult or made no difference had mid or lower proficiency scores, whereas the vast majority of those with higher proficiency scores responded that English spelling was easier. In a similar vein, the mean proficiency of those who found Arabic spelling easier was generally lower than those who found it difficult or did not make a difference, as was the case for those who found no written input easier. Therefore, there is some indication of the anticipated proficiency effect, where those with lower L2 English proficiency prefer shared Arabic OI or no OI, meanwhile those with higher English proficiency have a clear preference for English OI.

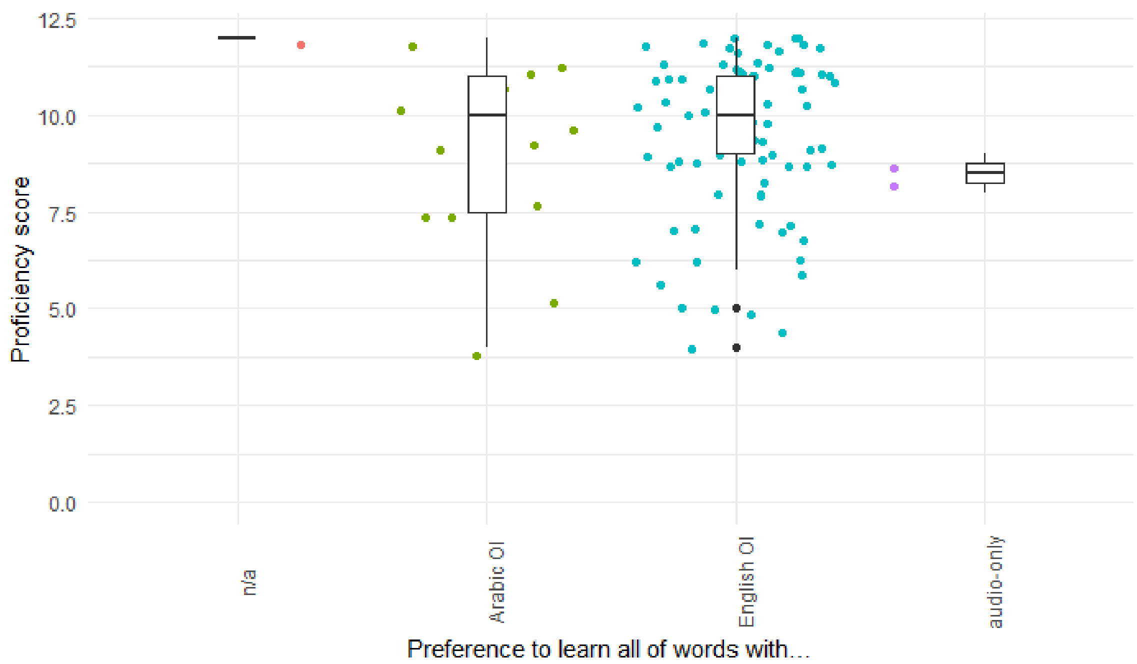


Figure 5.5: Questionnaire responses for overall spelling preference, by proficiency score

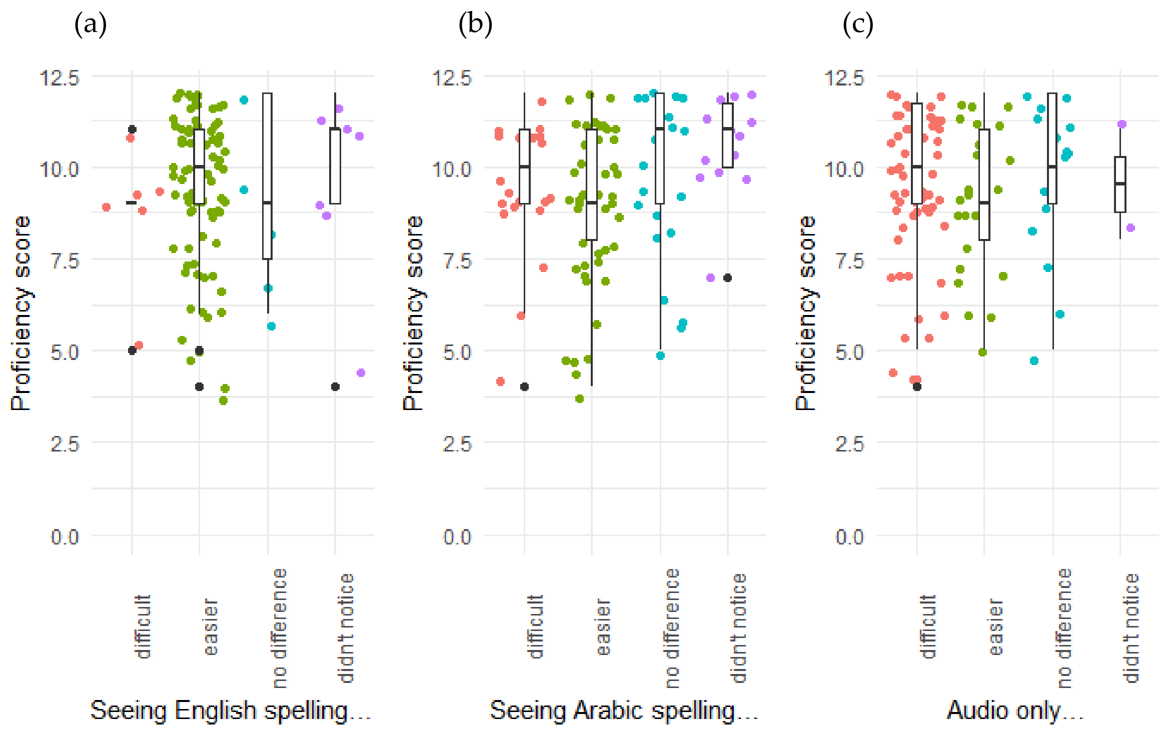


Figure 5.6: Questionnaire responses for perceived influence (a) English OI, (b) Arabic OI, and (c) no OI, by proficiency score

5.4.2 QCA of perceived OI influence

To further investigate participant perceptions about how OI influenced the way they learned the novel words, open responses to Q6 were analysed using QCA. The procedure for this analysis was laid out in section 5.3 and the final categories are presented in more detail here. Every unit of analysis was coded for general type of influence (positive, negative/no influence) and then further categorised according to details about the influence of written forms on the ability to process or remember the target items. The analysis resulted in five positive categories, including five subcategories, and three categories where OI was not perceived to be helpful. The positive responses all implied an enhancement of learning with written forms, and follow on from the phrase "OI helps…", to capture the facilitatory effects. The negative and neutral responses included more information regarding what it was about the written forms that made them unhelpful or unnecessary. Therefore these categories were amalgamated into one, which followed on from the phrase "OI does not help because it is…". The general types, main categories, and subcategories are visualised in figure 5.7.

The general types of responses are defined with examples in table 5.6, while the main and subcategories are defined with examples in table 5.7. These tables then formed part of the guidelines for coders in both languages (see appendix XI). Intercoder agreement was calculated using Krippendorff's alpha and resulted in unanimous agreement for the general response types, for both L1 English and Arabic coders, of either "OI helps" or "OI doesn't help" ($\alpha = 1$) . Agreement between L1 English coders was lower for the specific categories ($\alpha = .71$), but still above the threshold for acceptable agreement, especially as a exploratory study. Meanwhile, agreement between the L1 Arabic coders for the specific categories was high ($\alpha = .85$). Lower agreement was often related to subcategories within an agreed main category. For example, two coders opted for the subcategory *map sounds to letters*, whereas the other coder opted for the main category *connect audio and visual input*. Categories could be collapsed, prioritising the main category for arguably more accurate categorisation of the data (Mayring, 2014). However, as the acceptance level of 0.67 was surpassed (Krippendorff, 2004), agreement was satisfactory and detail was considered valuable.

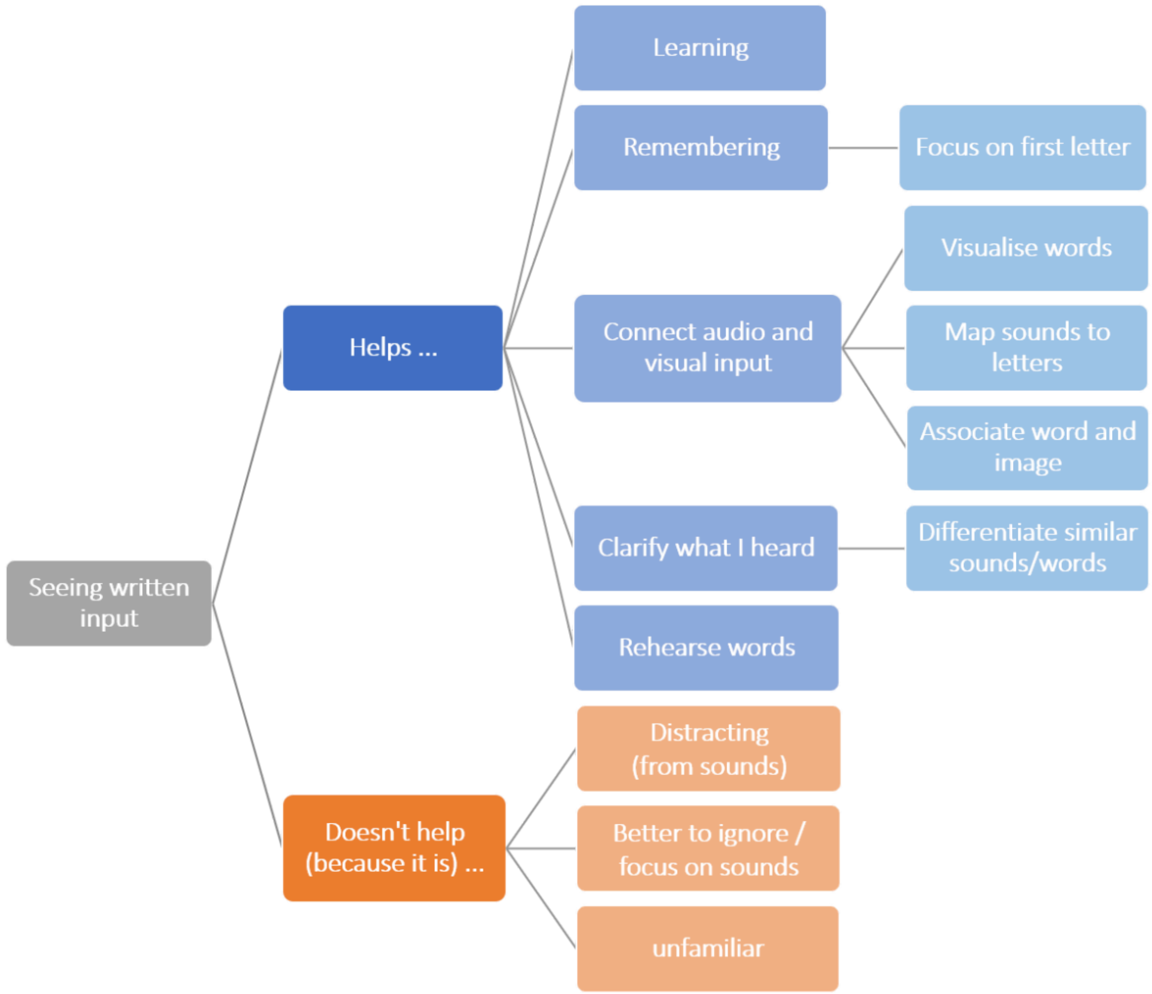


Figure 5.7: Hierarchy diagram showing the relationship between the general types of responses, the main categories, and subcategories for coding responses to Q6

Table 5.6: Coding guidelines for Q6 overall evaluation of OI influence during the task

Type	Definition	Examples
Positive influence (OI helps)	Seeing the written form was helpful, easier or facilitated learning, understanding or memory. This could be general or specific to English or Arabic script input **English keywords** easier/ faster/ help/ aid/ enable/ facilitate /able **Arabic keywords** أصبح/ جعلها أسهل/ أسرع, يساعدني [*more difficult/ makes it easier/ faster, helps me*]	"Easier to remember" "It's very helpful when i see the word written." "اذا كانت الكلمة مكتوبة بالانجليزية اسهل" [*If the word is written in English it's easier*]
Negative / Neutral influence (OI does not help)	Seeing the written form was unhelpful, introduced added difficulty or influence that inhibited learning, understanding or memory. Alternatively, seeing the spelling did not make a noticeable difference. Even if there was no explicit negative effect, it did not help. **English keywords**: difficult/ no help/ no difference/ distraction **Arabic keywords** لا يوجد فرق, كانت اصعب/ يشتت [*There's no difference, it was harder/ distracting*]	"Arabic written words were more of a distraction." "I read them when I noticed them but it didn't seem to help" "الكلمات المكتوبة كانت اصعب" [*The written words were harder*]

Table 5.7: Coding guidelines for Q6 specific perceptions of OI influence

Type	Category	Definition	Examples
OI helps…	Learning	Facilitates learning in general. Makes learning easier/ better/ faster. It does not specify further. **English keywords** learn(ing); easier/faster/ better **Arabic keywords** تعلمها/ استيعابها اسهل/ اسرع / أكثر [*learn it/ absorb it easier/ quicker/ more*]	"Yes, it makes learning easier." "التعلم اسرع بوجود الكلمات المكتوبة" [*learning is faster with the written words*]
	Remembering	Facilitates memory/ recall/ retention of words or objects. Does not specify beyond general memory/ memorisation aid. **English keywords** remember/ memory/ memorisation/ recall **Arabic keywords** يسهل التذكر / الحفظ [*easy to remember/ memorise*]	"Much easier to remember" "رؤية الكلمات العربية سهل تذكرها بالنسبة لي اكثر من الانجليزية" [*seeing the Arabic word makes it easier to remember it, in my opinion, more than in English*]
	Focusing on first letter	Specific to initial letter and not sound; use of first letter as a memory aid. **English keywords** First letter; begin/start with; remember	"I think it was easier to remember the first letter"

	Arabic keywords تذكر الحرف الأول [*remember the first letter*]	
Connect audio and visual input (broadly)	Useful to connect the different types or sources of input; references ideas that align with multimodal input and dual-coding; does not specify whether association is between spelling, sound or image. May reference learning and memory. **English keywords** connect/ associate/ match/ link/ together; hearing/ audio/ sound and seeing/ visual/ image/ word **Arabic keywords** ربط البصر والصوت (معًا/ سويا) [*connect sight and sound (together/both)*]	"Seeing English spelling helped to make connections between what I was hearing and what I was seeing." "الاستماع والقراءة معا تجعل من التعلم أسهل" [*listening and reading together make learning easier*]
Visualise words	Specifically mentions visualising the word in the mind; implies a focus on the whole word and does not mention the image or meaning. **English keywords** visualise/ picture/ imagine **Arabic keywords** تخيل شكل الكلمة [*imagine shape of the word*]	"I could visualise what I was hearing" "نعم فقد ساعدني في تذكر شكل الكلمة كثيرا سواء الكلمة بالعربية او الانجليزية فأنا أتخيلها" [*yes, it really helped me remember the word's shape, both the word in Arabic or English, so I imagined it*]

Map sounds to letters	Specifically mentions the connection between spelling and sounds; focus on individual letters and sounds rather than the whole word; may reference active processes of spelling out or reading along with sound. **English keywords** map/ connect letters and sounds; decode; spell **Arabic keywords** ربط/ تطابق الحروف والصوت [*connect/ match letters and sound*]	"In English it helped because it gave me something else to remember and map sounds onto." "رؤية الكلمة الإنجليزية يساعدني أحيانا في استيعاب المسموع وخاصة أن الكلمات الإنجليزية تطابق حروفها الصوت المسموع" [*seeing the word in English sometimes helps me perceive/ absorb the audio, especially as the English words match their letters and audible sounds*]
Associate word and image	Specifically mentions the connection between the (spelling of the) word and the image. **English keywords** connect/ link/ associate image/object/picture and sound/ word **Arabic keywords** ربط/ تطابق الصورة والصوت/ الكلمة [*connect/ match picture and sound/word*]	"The English words helped my brain link the images to the sound better." "كانت الكلمات المكتوبة باللغة الإنجليزية أدعى لتذكرها والربط بينها وبين الصورة" [*The written words in English were easier to remember and to connect them with the picture*]
Clarify what I heard	Clearly comprehend spoken form; check hearing was correct; identify sounds missed with audio only. **English keywords** know/ understand/ comprehend correct spoken form/ pronunciation	"Yes, in English because it helped me to clarify if I had heard the spoken pronunciation correctly." "الكلمات المكتوبة مصدر تأكيد للكلمة المسموعة" [*the written word is a source to confirm the spoken word*]

		Arabic keywords تأكيد/ معرفة للكلمة المسموعة/ النطق الصحيح, لفهم أفضل [*check/ know the audible word/ correct pronunciation, understand better*]	
	Differentiate similar sounds/words	Distinguish words and sounds that are similar. May mention the first letter in the context of differentiating similar sounds, rather than as a memory aid. **English keywords** distinguish/ differentiate/ separate; similar sounding **Arabic keywords** التفريق/ التمييز بين الكلمات/ الأصوات المتشابهة [*differentiate/ distinguish between similar words/ sounds*]	"I think it aided the distinction between similar sounding words" "الفرق بدى واضحًا في الكلمات المتشابهه مثل Fan Van " [*the difference is clearer in similar words like Fan Van*]
	Rehearse word	Practise the word out loud or in their mind. **English keywords** practice/ rehearse/ repeat; out loud; in my head/ mind **Arabic keywords** تدرب/ كرر على الكلمات/ الأصوات [*practice/ repeat the words/ sounds*]	"Seeing them helped me rehearse the word in my head."
OI doesn't help because it is…	Distracting (from sounds)	Written words were a distraction in general; related to cognitive overload or inconsistent mappings between sound and spelling.	"The Arabic words put with the picture made it so that I felt there

	English keywords Distract, too much **Arabic keywords** يشتت [*distract*]	was more to take in, so distracted from learning the sound” “في الغالب لا يحدث فرق بل ربما يشتت قليلا النظر الى الكلمة المكتوبة والصورة والاستماع" [*often, it didn’t make a difference, it probably distracts a bit looking at the written word, picture and listening*]
Unfamiliar	Written forms were written in an unfamiliar script, so unable to read or distinguish the words. **English keywords** Unfamiliar; couldn’t read; don’t know/ understand **Arabic keywords** غير مألوفة, لا أستطيع قراءتها [*unfamiliar, not able to read it*]	“The Arabic spelling didn’t make any difference because I do not know how to read it.”
Better to ignore/ focus on sounds	Ignored written forms; chose to focus on the sound and/or image **English keywords** Ignore; focus/ concentrate on/ pay attention to sounds; **Arabic keywords** أركز على الصوت, كان سماع الكلمات أفضل/أسهل [*focus on the sounds, listening to the words is better/ easier*]	“I focus more on hearing them out loud and trying to match the sound to the colours and shapes.” “لا، أركز فقط على الصوت" [*no, I focus only on the sounds*]

The full data set, including the tags with the final categories can be found in appendix XIV for both L1 English and Arabic-speaker responses. This demonstrates how responses were coded, as well as showing the responses that were not deemed sufficiently specific or relevant to be coded. Overall, 37 L1 Arabic responses were excluded, on the basis that responses were insufficiently specific to code, while only 4 L1 English responses were excluded for this reason. Such answers included, “yes, a lot”, “70%”, “just a little”. These responses indicate an awareness of orthographic influence, however the direction of that influence was not clear. Further, the present study is focused on the *type* of influence, rather than the *amount*. While the answer “yes” could not be coded, as this could be interpreted as “yes - there was a positive influence” or “yes - there was a negative influence”, a “no” response implied a neutral influence of OI. Therefore, short answers indicating a perceived lack of influence were coded as “OI doesn’t help”, but were not coded with more specific categories.

The rest of the analysis in section 5.4.2.1 and 5.4.2.2 focuses exclusively on the responses that contained enough information to be coded for both general type and specific categories of influence. Bearing in mind that the unit of analysis was each clear semantic unit, longer responses could be broken down into multiple units of analysis, each tagged with a category. Therefore, the combination of fewer exclusions and longer responses, containing multiple units to code, meant that there were more units of analysis for the L1 English group than the L1 Arabic group. The results reported below represent the data from 116 L1 English units but only 59 L1 Arabic units, with an overview provided in table 5.8.

Table 5.8: Overview of Q6 QCA results with number of analysis units for each category (% of total units), by L1 group

Strategy category	Strategy subcategory	L1 Arabic	L1 English
OI helps…			
Remembering		18 (31%)	23 (20%)
	Focus on first letter	*0 (0%)*	*8 (7%)*
Connect audio and visual input		16 (27%)	28 (24%)
	Associate word and image	*6 (10%)*	*11 (10%)*
	Map letters to sounds	*4 (7%)*	*7 (6%)*
	Visualise words	*2 (3%)*	*4 (3%)*
Clarify what I heard		8 (14%)	33 (28%)

	Differentiate similar sounds	*6 (10%)*	*20 (17%)*
Learning		12 (20%)	2 (2%)
Rehearsing		0 (0%)	5 (4%)
OI does not help…			
Better to ignore/focus on sounds		3 (5%)	16 (14%)
Distracting (from sounds)		2 (3%)	4 (3%)
Unfamiliar		0 (0%)	5 (4%)

5.4.2.1 Written input helps

When asked whether seeing written input made a difference to learning the novel words in the study, 54 (92%) L1 Arabic units and 91 (78%) L1 English units indicate that OI was perceived to help learning. Within those units, there are 19 examples where L1 English-speakers specify that English spelling is helpful, in comparison to five examples from the L1 Arabic-speakers. Only two Arabic-speakers state that seeing the Arabic spelling is easier than in English. There is an additional example where one Arabic-speaker reports that it does not matter whether the word is written in English or Arabic script, both are helpful. Meanwhile a different participant makes the point that both could be useful, if the Arabic spelling includes vowel diacritics (see 2.3.1).

رؤية الكلمة الإنجليزية يساعدني أحيانا في استيعاب المسموع وخاصة أن الكلمات الإنجليزية تطابق حروفها الصوت المسموع. قد تؤدي الكلمات العربية الدور نفسه لو كانت الكلمات العربية مشكلة بالحركات (AR5370516)

[*Seeing the word in English sometimes helps me to perceive/absorb the audio, especially as the letters in the English words match the audible sounds. Arabic words may play the same role if the Arabic word forms are "vowelled"*].

This response demonstrates an awareness of the systematicity of grapheme-phoneme correspondences, indicating that overt mappings between sounds and letters are critical to OI being useful. It also notes the varying transparency of the Arabic script, depending on whether the diacritics marking short vowels are present. Additional insights into how participants conceptualised the positive influence of OI are investigated by examining the specific categories described in section 5.4.2 and how they align with participant responses. Categories are presented in order of the number of L1 Arabic units within each category[33].

Remembering

The role of written input in supporting memory is mentioned in 18 L1 Arabic and 23 L1 English response units. Many responses refer to the positive influence of OI on remembering new words across different categories. However, the units of analysis that have been attributed this category are broad and offer little information beyond remembering. Some specified that remembering was easier or faster with written input, or that it acted as a "trigger" for memory. While often general in their content, a subcategory was formed in relation to responses that specified a *focus on the first* letter as a memory device.

> Easier when they were written in English - I think it was easier to remember the first letter (EN5758767).

There were no examples of this within the L1 Arabic data, whereas there were eight units that mention this in the L1 English data. These responses focused on English OI and mentioned both distinguishing and remembering the first letter of the new words. As the first sound was the target contrast of the minimal pairs, it is difficult to say whether focusing on the first letter was related to disambiguating the similar sounding pairs or a more general memory tactic. However, as is described in the following categories, other responses appear to target differentiating sounds specifically, whereas these responses focus more on remembering letters rather than sounds.

Connect audio and visual input

The category with the largest total response units is *connect audio and visual input*, with 16 L1 Arabic and 28 L1 English units. All examples refer to written forms facilitating a connection between the audio and visual information, often suggesting this supports the target item in memory. One L1 English-speaker explicitly mentions "dual-coding", referring to the theory that there are associative mental structures, networks of imagery, and verbal

[33] The category *learning* is presented at the end, even though it was not the smallest category, as most responses were vague and warrant little description.

representations, which can be activated and developed to support learning (Clark & Paivio, 1991). This idea of learning and remembering through associating the visual and verbal is reflected in other responses, such as below:

> Yes ,i prefer the word written while listening to it , that's how i memorize the word twice written and spoken (AR5015277).

Similar ideas are reflected in responses within the three subcategories, which specify that OI helps associate the new word with the corresponding image, map sounds to letters and visualise the words. There are six L1 Arabic and 11 L1 English responses which *associate the word with the image*. Some L1 English responses focus on a general mental association between the word and the object (EN5328002), whereas others emphasise unifying the shape of the spelled word with the correct image, as a joint picture in the mind (EN5338550). Both views are echoed in the L1 Arabic responses.

> Watching the written words make learning easier as I can link the image with the written word (AR5031281).
>
> Yes, easier to associate the word with the object when you can see the spelling (EN5328002).
>
> I felt it was easier to remember the picture and the word together when they flashed up on scren together - like trying to take a mental screenshot (EN5338550).

Rather than focus on the connection between the written word and the pictured meaning, four L1 Arabic and seven L1 English responses report that written input helps *map letters to sounds*. Of these, two L1 Arabic and four L1 English responses make specific reference to English spelling. L1 English responses mention breaking words down into sublexical units so that it is possible to "spell it out in my head" (EN5308960) or decode the word as they were hearing it. Others stated that this was an additional memory aid or that they would parse the sounds through the pronunciation they associated with the shared letters. The example below shows the conflation of letters and sounds, where the participant indicates that the letters correspond to fixed pronunciations, which are used to anticipate and assess incoming speech.

> It helped to see the English versions written out to match sounds of pronunciation with what I was hearing. (EN5649032)

The relationship between sounds and letters is also explored in the L1 Arabic responses, such as the one mentioned above by AR5370516, in reference to the more transparent relationship between the English spelling and the audio, compared to the unvowelled Arabic spelling. Few would argue that English transparently denotes vowels in spelling, but this was intentional in the design of the pseudowords, so as not to create confounding difficulty alongside the target consonantal contrasts. The relevance of orthographic depth to the influence of written input is raised by another participant, but from quite a different perspective. Interestingly, they argue that English written forms are helpful *because* English orthography is generally inconsistent, acknowledging that the written form cannot easily be deduced from the spoken form. This implies that the participant is thinking beyond the task, drawing on existing knowledge of English orthography and phonology and considering the need to use new vocabulary in both spoken and written modalities.

> For English written forms are helpful as English is not a consistent language (written forms to spoken forms) (AR5196532).

The final subcategory within the main category of *connect audio and visual input* refers to how written input helps *visualise the words*. Two L1 Arabic and Four L1 English-speakers mention visualising, picturing, or imagining the word in their mind. Some examples specify that this process is useful for remembering the word later, with one noting that visualising the word in this way can include the whole word or at least one letter.

> نعم فقد ساعدني في تذكر شكل الكلمة كثيرا سواء الكلمة بالعربية او الانجليزية فأنا أتخيلها واتذكر لو حتى حرفا واحدا
> (AR5357676)

[*Yes, it helped me a lot with remembering the shape of the word, whether the word was in Arabic or English, so I imagine and remember, even if only one letter.*]

Clarify what I heard

The next largest category emphasises the role of OI in clarifying what you hear. There are eight L1 Arabic and 33 L1 English units that are included in this category, where six L1 Arabic and 20 L1 English units are placed in the subcategory of *differentiate between similar sounds/ words*. Those that are included in the main category mentioned being more sure about the spoken form or knowing the correct pronunciation, as well as drawing attention to mishearing. A mistrust of audio information alone and perceiving the written form to be more correct or even "objective", as EN5750994 says, is found in both language groups. There are also further examples of the blurred lines between sounds and letters, the visual and the verbal, and spelling compared to pronunciation.

نعم ، لمدى كثير جدا بالمكتوبه استطيع معرفه النطق الصحيح (AR5082295)

[*Yes, in a really big way, in writing I was able to know the correct pronunciation*]

Provided context to the pronunciation and provided more objective concepts in my head of what the thing was than audio alone (EN5750994)

It makes a difference to me in general because I can see the sounds I am unable to pick up by just listening. (EN5312160)

Within the subcategory *differentiate similar sounds/ words*, several participants demonstrate noticing the minimal pairing between similar sounding words, focusing on *f-v* and *m-n*. As mentioned, some responses lack clarity as to whether distinguishing similar sounding words and their initial sounds was to aid the remembering of words by their first letter, or to improve understanding. However, all the examples in this subcategory focus on differentiating similar sounds with no reference to memory, implying a distinct process centred more on comprehension than recall. It is notable that, while the L1 English responses mention both *f-v* and *m-n* initial sounds, the L1 Arabic responses focus on *f-v* alone without mention of *m-n*. This implies that, for these participants, written forms specifically supported the distinction between confusable contrasts. Some L1 English responses also mention other sounds or subtle differences, not just the target contrasts.

> نعم، أصبحت الفروق واضحة بين F و V. (AR5339892)
>
> [*yes, it makes the difference between V and F become clear*]
>
> Seeing english written words helped me differitiate between the different similar words,ie, helped notice if it ended in T rather S. (EN5338600)
>
> yes. accentuates the difference between f / v sounds, m / n sounds (EN5706490)

Learning and rehearsing new words

The next category to mention is the general facilitation of *learning*. There are 12 L1 Arabic and two L1 English responses included in this category. Most responses are relatively vague, with little specification of influence beyond being broadly helpful and making learning easier or faster. The final category, which is only reflected in the responses of L1 English-speakers, is the role of written input in facilitating the *rehearsal of words*. There are five L1 English responses that refer to using the written forms to practise saying words out loud, repeating them, or rehearsing the words in their mind, to help them remember the new words.

> I was also able to remember the spelling and repeat. hearing the word aloud from myself helped me (EN5676564)

Overall, evidence has been given for a variety of positive perceptions related to written input, when learning new words in the study. Across both language groups, participants perceive written input to facilitate the connection between audio and visual input, mentioning the association between words with their pictured meaning, mapping sounds to letters, and visualising all or part of the written word. These responses often include the positive impact that written input has on remembering new words, which is further highlighted in its own category. A large proportion of participants also note the role of OI in clarifying or confirming the spoken form, mentioning the visual distinction between similar sounds and words. The facilitative impact of written input was also connected to learning broadly. Two additional categories only included responses from the L1 English group, namely the subcategory *focus on the first letter*, within *remembering*, and finally the category of *rehearsing* the words. The lack of L1 Arabic responses within these categories does not mean that these ideas, behaviours, or processes were not present in the participants, simply that they were not reported in response to this question. In fact, related ideas are reflected in L1 Arabic responses analysed in section 5.5 and chapter 6.

5.4.2.2 Written input does not help

In contrast to the positive perceptions of orthographic influence in the previous section, only five (9%) L1 Arabic and 25 (22%) L1 English response units indicate that written forms did not help learn the new words, either because it made no noticeable difference or introduced additional difficulty. Within those responses there are eight examples where L1 English participants specify that Arabic spelling did not help. There are also three L1 Arabic responses that mention that the Arabic spelling was particularly difficult[34]. None of the L1 Arabic responses mention English spelling, whereas four L1 English speakers reference English spelling. The negative and neutral perceptions of orthographic influence is explored within the following subsections, where OI does not help because it is: *better to ignore/ focus on sounds*, *distracting (from sounds)*, or *unfamiliar*.

Better to ignore / focus on sounds

Most responses suggesting that written input is not helpful emphasise the importance of focusing on sound, as well as noting both conscious and unconscious efforts to ignore the written form. Three L1 Arabic and 16 L1 English units are included in this category. Within the L1 English responses, participants report ignoring one of the scripts or a preference for focusing on the sound alone. Some responses indicate a conscious suppression of attention to spellings, whereas others mention a subtle adjustment of attention that perhaps they only recognised upon reflection[35]. The issue of transparency and incongruence between spelling and pronunciation is also raised here (EN5521508).

[34] Two of which were given the general code of "OI does not help" but no specific code (AR5336601; AR5499647), therefore falling outside the main analysis but can be viewed in appendix XIV.
[35] This category also includes the participant who was previously mentioned not recalling seeing any written input during word learning (EN5588274).

> No difference. hearing words make it easier than reading them, especially in Arabic. (AR5010775)
>
> I preferred learning from sound alone, as English spelling didn't reflect pronunciation (EN5521508).
>
> In Arabic, I just ignored it so it made no difference (EN5663887).

Distracting (from sounds)

Orthography is also reported to have a distracting influence, referenced in the responses of two L1 Arabic and four L1 English participants. Across these responses, participants suggest that written input distracts from sounds or the picture, partly because of cognitive overload. Trying to take in the new word, its spelling, and a novel image at the same time is too difficult. Within the L1 English responses, some mention script specific influences, such as finding Arabic spelling distracting. Meanwhile, another person mentions that the visual information within the English spelling distracts from the sound of the word. Overall, these participants implied that the multimodality of input was too much to attend to, in contrast to those who felt it supported memory through dual-coding.

> The English words I understood so it didn't make too much difference but the Arabic words put with the picture made it so that I felt there was more to take in, so distracted from learning the sound and I was trying to correlate th sound to the shape of the word?!! (EN5595978)

Unfamiliar

While not mentioning it explicitly, the example above (EN5595978) implies that the distraction of Arabic spelling is related to the unfamiliarity of the script. This is expressed more overtly in the responses within the *unfamiliar* category. There are five L1 English responses within this category and no L1 Arabic responses, which is logical as neither the English or Arabic script were unfamiliar to the L1 Arabic group. These responses all reference not understanding, being able to read or distinguish between words that were written in the Arabic script. In general, this is expressed to have a neutral rather than negative impact, as most state that unfamiliar OI makes no noticeable difference to learning.

> The Arabic spelling didn't make any difference because I do not know how to read it, the only thing I did was trying to see a similarity between the shape of the words and the sound of them (EN5765136).

In summary, a minority of participants in both groups perceived that written input did not help them learn the new words. This was noted by a larger proportion of the L1 English group compared to the L1 Arabic group. However, across both groups the dominant category was the perception that it was *better to ignore written input/focus on sounds*. This could be related to the other categories, where both L1 English and L1 Arabic participants mentioned that written input could be distracting or overloading when presented with new word forms and meanings, as well as issues related to the unfamiliarity of Arabic spelling for L1 English participants.

5.4.3 Summary of perceived OI influence

Responses to the multiple-choice questions 1-5 (section 5.4.1) revealed that most participants perceived written input to be supportive when learning new words in the experiment, with both L1 English and L1 Arabic-speakers indicating a preference for learning the words with English spellings. A small majority of the L1 Arabic group indicated that Arabic spelling also had a facilitative effect, while the L1 English group mostly perceived unfamiliar Arabic OI to make no difference. Of the three different written input conditions, participants generally perceived learning to be more difficult without any OI. However, this was more pronounced in the L1 Arabic group.

These group tendencies were further illuminated by exploring the open responses to Q6. Figures 5.8 and 5.9 display the proportion of response units that were included within each identified category and subcategory, by L1 group. These treemaps indicate that, while both groups perceive OI to be predominantly positive, this was more evident in the L1 Arabic responses in comparison to the L1 English responses. Although, it is noteworthy that there were around double the number of response units from the L1 English group. While frequencies and proportions differ, there is evidence across both groups that written input is seen to help connect audio and visual input, especially in relation to associating the word with its corresponding picture and mapping sounds and letters. Participants mention ideas related to dual-coding as advantages for memory, in particular.

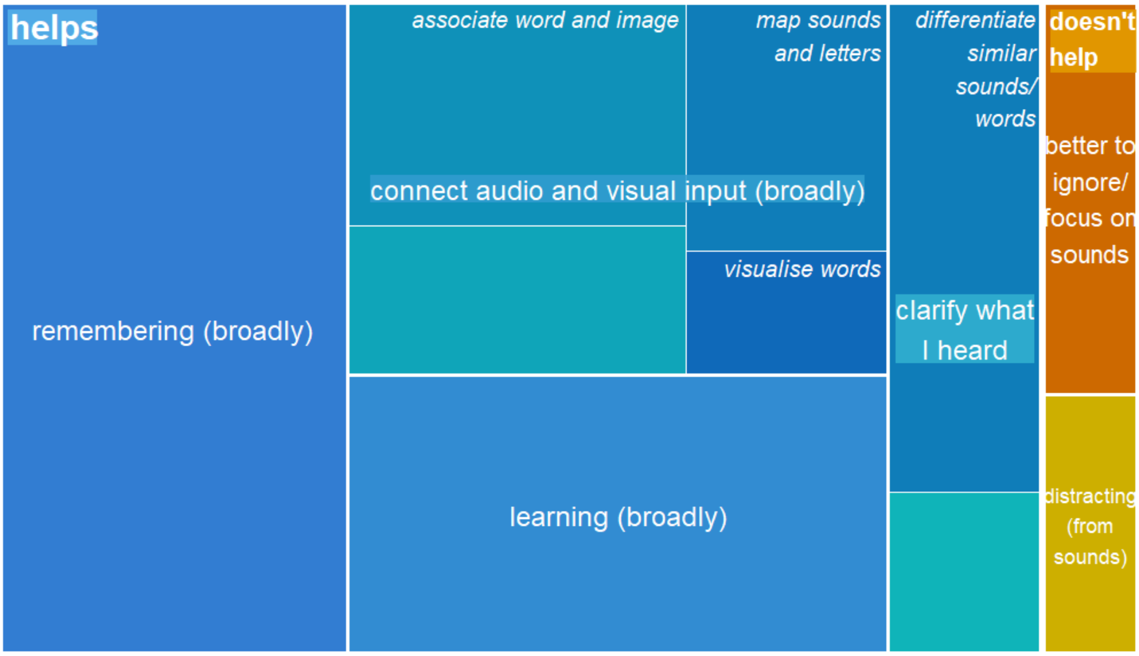


Figure 5.8: Treemap of L1 Arabic group's perceived influence of OI during the study

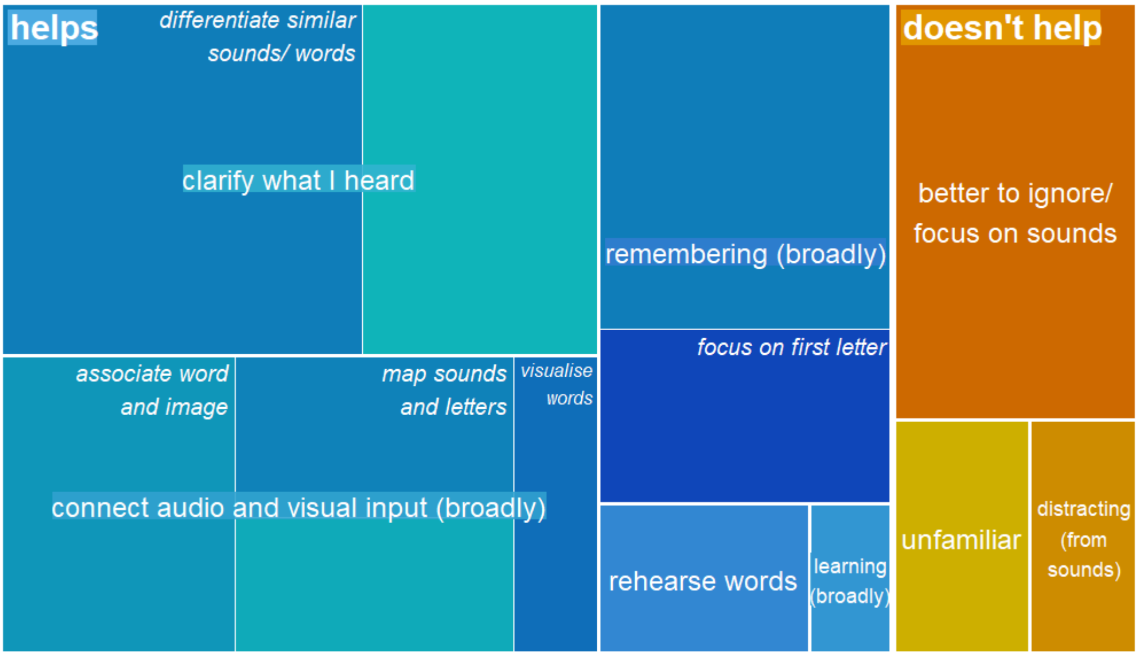


Figure 5.9: Treemap of L1 English group's perceived influence of OI during the study

The positive influence of written input on remembering the new words is noted broadly, with some L1 English-speakers focusing on the first letter of new words as a memory aid. Others focus on the first letter and initial sound to clarify their comprehension of new words, especially those that sounded similar. Meanwhile, some perceive written forms to help with more general clarification of spoken forms, to avoid mishearing and provide insight into the "correct" form of the word. Within the L1 Arabic group particularly, but not exclusively, participants referred to a broadly positive impact on learning as a result of seeing the written forms. A small number of the L1 English group also mention the use of written forms when wanting to rehearse or repeat the new words to themselves.

In contrast, responses from both language groups, but more so the L1 English group, note the possibility that written input could have a null or negative influence on learning. The majority of those who did not perceive written input to be helpful refer to ignoring spellings or focussing on sounds, either consciously or somewhat unconsciously. Responses from both groups suggest that written input can be distracting or introduce cognitive load issues that make it difficult to fully attend to the sound of words. Finally, a small number of L1 English-speakers draw attention to their unfamiliarity with the Arabic script. However, this is generally expressed to have a neutral rather than negative impact on learning.

In relation to the research question stated at the start of this chapter, most participants across both languages perceived orthographic input to be influential during the study. This influence was predominantly positive with a general preference for viewing the English spellings, suggesting that participants perceived a distinct influence between OI in different scripts. However, it is of interest that the difference between English and Arabic OI and the reported benefits associated with written input, particularly differentiating confusable contrasts, were not apparent in the matching task results. Overall, the articulated perceptions related to orthographic influence were complex and varied across individuals, yet there was surprising unity across the two language groups. Within both groups, there was reference to established concepts, such as orthographic depth, congruence, familiarity and perceptibility, as well as theories of learning and memory. The next section presents the analysis of Q7 and Q8, to shed further light on these findings and add context from broader beliefs related to the importance of written input .

5.5 Participant beliefs about the importance of written input

As with the previous section, analysis is exploratory and there are no explicit hypotheses, due to the dearth of qualitative research into learners' beliefs around the importance of written input when learning a second language, especially across writing systems. Based on anecdotal conversations and observations as a language teacher and learner, it is anticipated that participants from both language groups will place a high degree of importance on seeing the written form when learning new words, in general. This aligns with findings reported above in section 5.4, where most participants perceived OI to have a positive influence on their learning during the tasks in the study. I expect that these perceptions will extend to more general beliefs about the importance of written input when learning new words in a language. It is also expected that this view will align with previous learning experiences. The first relevant item on the questionnaire, Q7, asked for a binary yes-no response as to whether it was generally important to see new words written down the first time you hear them, when you are learning a language. This was followed by the open Q8, to deepen understanding of why they held a particular belief.

5.5.1 Frequencies of perceived OI importance

As predicted, participants in both language groups reported that it was important to see the spelling of a new word when learning a language. Figure 5.10 demonstrates that the proportions of responses differed between the two groups, where 100 (88%) L1 Arabic-speakers indicated “yes” it was important to see the spelling when learning new words, whereas 84 (72%) L1 English-speakers gave this answer. Thus, the importance of written input was more strongly reflected in the L1 Arabic group, while a larger proportion of L1 English-speakers reported that it was not necessarily important to see the written forms when learning new words.

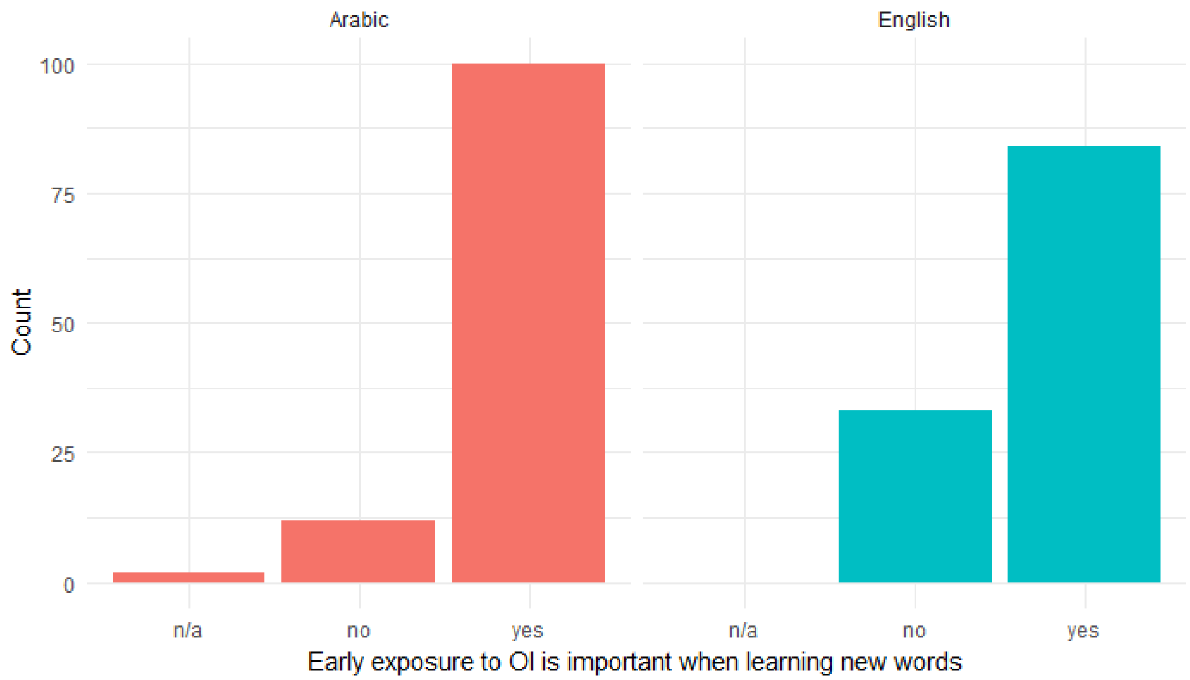


Figure 5.10: Questionnaire responses for Q7, by L1 group

5.5.2 QCA of perceived OI importance

To further investigate participant beliefs about the importance of written input, open responses to Q8 were analysed using QCA. The procedure for this analysis was laid out in section 5.3 and the final categories are presented in more detail here. Every unit of analysis was separated into broad types by the yes-no response to Q7, and then categorised according to justifications about why they believed OI to be important or not. Attention was paid to the ways participants believe OI facilitates, inhibits, or proves redundant when learning new words. The affirmative responses follow on from the phrase "early OI is important for…", whereas the negative responses follow on from "early OI is not important because…". The main and subcategories are defined with examples in table 5.8 and a hierarchical diagram of the categories is also displayed in figure 5.11. Intercoder reliability was calculated for L1 English and L1 Arabic data separately and both groups of coders had identical levels of agreement (α = .76)[36]. Templates of the coder guidelines can be found in appendix XII, and the R script for calculating intercoder reliability is in appendix XIII.

[36] While a high level of agreement would typically be >.8, >.7 is considered acceptable, especially for more exploratory content analysis (Gläser-Zikuda et al., 2020; Mayring, 2014).

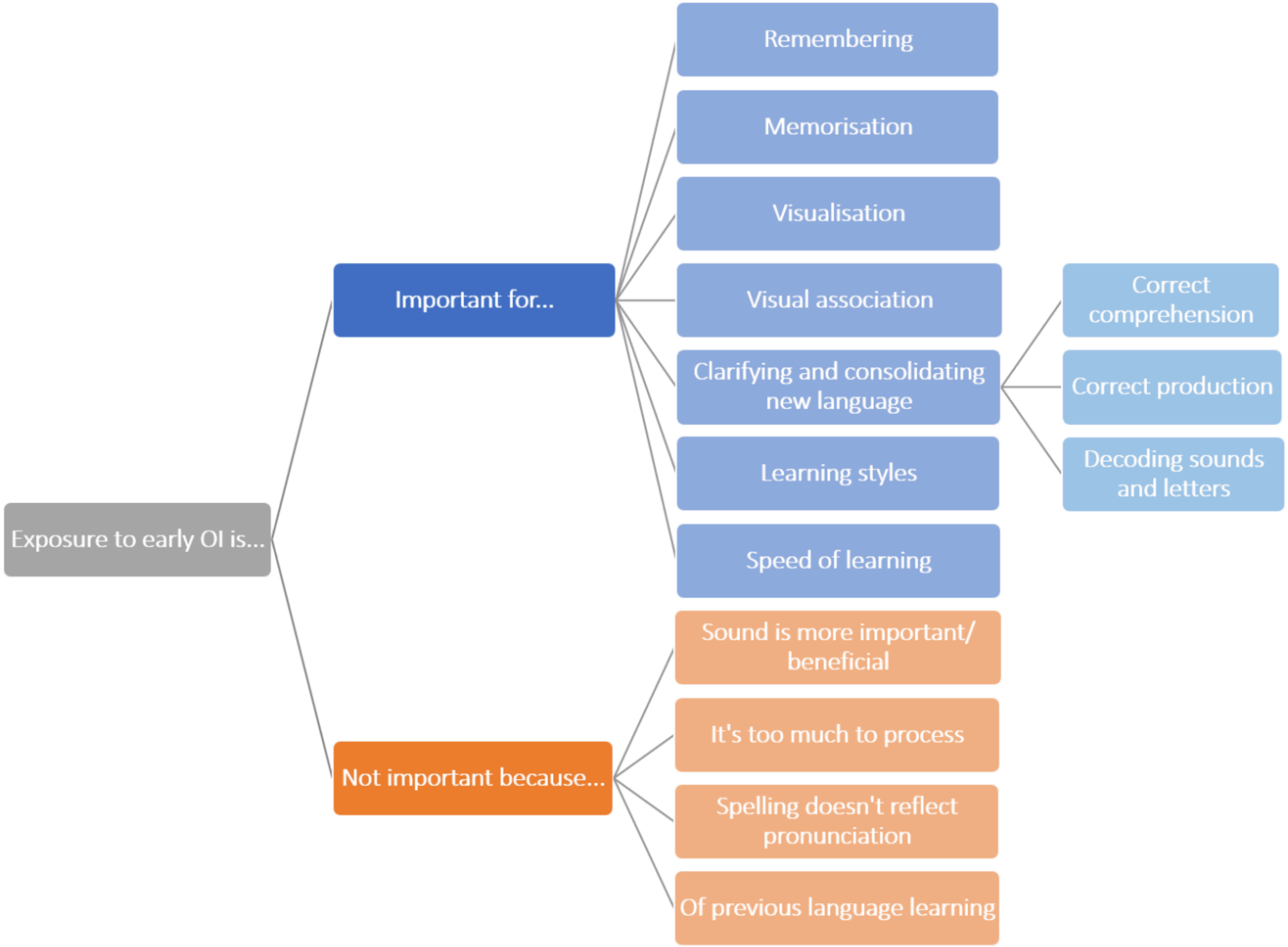


Figure 5.11: Hierarchy diagram showing the relationship between the general types of responses, the main categories and subcategories for coding responses to Q8

Table 5.9: Coding guidelines for Q8 beliefs about OI importance

Type	Category	Definition	Examples
Early OI is important for…	Remembering (broadly)	The additional visual cue from written input helps remember the words, in general. Does not specify further. **English keywords** Remember, memory, recall **Arabic keywords** تذكر, ذاكرة [*to remember/ memory*]	"I think it helps me remember them" "Additional aid to memory" "يسهل تذكرها" [*it's easier to remember it*]
	Visualisation	Visualising the (shape of the) word is helpful and may refer to memory and learning. Does not specify further. **English keywords** Visualise/ imagine/ picture in my mind/ brain **Arabic keywords** تخيل/ تذكر شكل الكلمة, في مخيلتي/ ذهني [*imagining/ remember shape of the word, in my imagination/ mind*]	"Helps me to see the word and visualise it." "حتى اتخيل شكل الكلمة" [*so I can imagine the shape of the word*]
	Visual association	Written input facilitates connections between verbal and visual input, as well as linking to existing linguistic knowledge. Not specific to individual sounds and letters, but broader connections between word, spelling and meaning. May contain ideas related to dual-coding, as well as memory and learning.	"You can make stronger connections between the word, its spelling and the sound and the concept it represents."

	English keywords Connect/ associate/ link the written word with sounds or meaning; together **Arabic keywords** ربط الكلمة المكتوبة مع الصوت أو المعنى / النطق و الكتابة معا [*connect the written word with sound or meaning/ the pronunciation and the writing together*]	"لكي اربط في ذهني الكلمة المكتوبة مع الصوت ... احفظ النطق و الكتابة معا في نفس الوقت" [*In order to associate the written word and sound in my mind … I memorise the pronunciation and the writing together at the same time*]
Memorisation	Connect the written input to the process of memorisation and rote learning. May make it a better process. Does not specify further **English keywords** Memorise, learn by heart **Arabic keywords** الحفظ [*memorise*]	"In order to memorise that word." "لكي يسهل حفظها" [*so it's easy to memorise*]
Clarifying and consolidating new language	Written input helps to "know" and feel sure about the correct form of the word. It helps recognize and embed new language in the mind **English keywords** Know / recognise / familiar/ embed / anchor/ consolidate in brain **Arabic keywords** تثبيت / ترسيخه/ ترتكز في المخ, معرفتها [*absorb/ establish/ anchor/ cement in the brain, know it*]	"In order to create a clear mental representation in my mind. "هو بمثابة تثبيت للكلمة" [*It serves to cement/establish the word*]

Correct comprehension	Specifically helps comprehend and confirm the correct form of the word. Avoid mishearing spoken forms, especially difficult to perceive sounds. Focus on perception and receptive language processing, not production. **English keywords** Mishear/ listen, understand/ comprehend, be sure/ confirm, differentiate/ discriminate **Arabic keywords** لتمييز/ تفريق بين كلمة أخرى شبيهة, لافهم, للتأكد ما استمعت إليه [*to distinguish/ differentiate between similar words, to understand, to confirm what you heard*]	"It prevents you from learning it incorrectly incase you have misheard the pronunciation" "للتأكد من صحة ما استمعت إليه" [*to verify what you heard*]
Correct production	Specifically helps pronounce or write a word or sound correctly. Avoids frustrations or insecurity around using the language accurately. Focus on usage and productive language processing, not perception. **English keywords** Pronounce/ say/ use correctly, speaking, writing **Arabic keywords** النطق الصحيح, استخدامها في التحدث والكتابة [*correct pronunciation, use it in speaking and writing*]	"You can also make sure you are pronouncing it correctly if you see it." "لأعرف نطق الكلمة." [*to know the word's pronunciation*] "لمعرفة كيفية كتابتها" [*to know how to write them*]
Decoding sounds and letters	Looking for patterns between sound and spelling, and breaking words down into smaller units. May mention learning phonetically, supporting reading, or vowels and diacritics in Arabic script.	"It helps for a learner to recognise patterns within the spelling system and

	English keywords Finding/ seeking spelling patterns, breaking down into sounds, phonetic, sounding out words **Arabic keywords** املاء وصوتها, الأنماط بين الحروف والأصوات, حركات التشكيل [*to know how to write them*]	how they interact with the sounds. "حتى اربط بين املاء الكلمة وصوتها" [*to know how to write them*]
Learning style	Identify their personal learning style, particularly being visual or textual learner. May state this as a preference. **English keywords** Learning style, visual, textual **Arabic keywords** ذاكرتي صورية, متعلم بصري [*visual memory, visual learner*]	"I am a more visual learner than an auditory learner." "لأني بصرية" [*because I am a visual learner*]
Speed of learning	Written input speeds up the process of learning, remembering, memorising, absorbing the new words. **English keywords** learn/ remember faster **Arabic keywords** حفظ/ تعلم بشكل أسرع, بسرعة [*memorise/ learn faster, quickly*]	"I can recognise it and remember it faster." "تثبت الكلمة اسرع" [*to absorb/establish the word faster*]

Early OI is not important (because…)	Sound is more important/ beneficial	Places importance on knowing the sound(s) well. This is the best point of focus. It helps to know and remember the phonological form of the word as well as rehearse it orally. **English keywords** Focus on sound, beneficial, important **Arabic keywords** التركيز على الصوت, الصوت أكثر أهمية [*focus on sound, sound is more important*]	"If the focus is purely on sound, hearing and then repeating orally is more beneficial in consistently replicating and identifying the sound/word."
	It is too much to process	Additional visual input is too much to process with the sounds as well. It interferes/ distracts from sounds, generally overloading. It introduces unnecessary elements that get in the way of learning. **English keywords** Overload, too much, interferes, distracts, unnecessary **Arabic keywords** اتشتت عن استماع, أكثر من اللازم [*distracts from listening, too much*]	"I think it interferes with my listening to see the word." "لانه يصعب علي المهمة في حفظ الاملاء للاحرف وحفظ الكلمة المستمعة" [*because it's hard for me to memorise the letters' spelling and memorise the heard word*]
	Spelling doesn't reflect pronunciation	Pronunciation does not consistently match the spelling of words. Spelling can be misleading. May refer to spelling-sound correspondences in specific languages (orthographic depth). **English keywords** Spelling doesn't match pronunciation, confusing/ misleading/ unreliable	"Spelling in English can be confusing." "Words can be pronounced differently to how they look."

	Arabic keywords الأنماط لا تتطابق مع النطق [*spelling doesn't match with pronunciation*]	“اللغة الإنجليزية أصوات غير متوفرة في حروف العربية” [*English sounds not available in Arabic letters*]
Previous language learning	Reference previous experiences of language learning, as evidence against reliance on OI. May include general ideas and observations about language learning or personal preferences and strategies. **English keywords** Prefer/ best way to learn, speaking and listening **Arabic keywords** أفضل أن أتعلم من خلال التحدث والاستماع, الطريقة المألوفة للتعلم [*I prefer to learn by speaking and listening, the familiar way to learn*]	“I prefer to learn languages by speaking and hearing them, not writing” “انا شخصية سمعية" [*I'm an auditory person*]

The full data set, including the tags with the final categories, is provided in appendix XV, for both L1 English and Arabic responses. Within these responses, there were several examples where participants answered the question in relation to learning within the experiment rather than language learning more generally, and were therefore excluded from this analysis. Overall, 14 L1 Arabic and 17 L1 English responses were excluded on the basis they were insufficiently specific to code or were made in reference to the experimental task. The rest of the analysis in sections 5.5.2.1 and 5.5.2.2 focuses exclusively on the responses that contained enough relevant information to be coded for the categories of why OI is believed to be important, or not. As with the previous section 5.4.2, longer responses could be broken down into multiple units of analysis, each tagged with a category. In contrast to the previous section, there was a more even balance between the units of analysis for the L1 English and Arabic groups. The results reported below represent the data from 116 L1 Arabic and 122 L1 English units, with an overview provided in Table 5.10.

Table 5.10: Overview of Q8 QCA results with number of analysis units for each category (% of total units), by L1 group

Strategy category	Strategy subcategory	L1 Arabic	L1 English
OI is important for…			
Clarifying and consolidating		47 (41%)	54 (44%)
	Correct comprehension	*15 (13%)*	*25 (21%)*
	Correct production	*16 (14%)*	*11 (9%)*
	Decoding sounds and letters	*6 (5%)*	*12 (10%)*
Remembering		17 (15%)	10 (8%)
Memorising		16 (14%)	0 (0%)
Visual association		11 (10%)	12 (10%)
Visualisation		7 (6%)	8 (7%)
Speed of learning		6 (5%)	0 (0%)
Learning styles		3 (3%)	9 (7%)
OI is not important because…			
It's too much to process		4 (3%)	7 (6%)
Sound is more important/beneficial		2 (2%)	10 (8%)
Spelling does not reflect pronunciation		2 (2%)	8 (7%)
Previous (language) learning experience		1 (1%)	4 (3%)

5.5.2.1 Written input is important

When asked whether or not it is important to see the written form when learning a new word, 107 (92%) L1 Arabic units and 93 (76%) L1 English units indicate that it is, indeed, important to see written input. In comparison to section 5.4.2, there is little direct reference to the comparative influence of English and Arabic orthography. However, within these responses, there are two examples where L1 Arabic-speakers reveal contradictory reasons for the importance of seeing English written forms, echoing different participants in section 5.4.2.1. One notes that English spelling offers more transparent information about the sounds of the word, such as vowels and consonants that are not reflected in Arabic spelling (e.g., /v/). In contrast, another (AR5247103) highlights the lack of consistency between sounds and spelling in English, and suggests this difference is why it is important to learn the spelling, presumably because the relationship is not otherwise intuitive. So, one believes that seeing English OI is important because sound-letters mappings are a transparent guide, while the other suggests that seeing the spelling is important precisely because it is opaque. These examples show an awareness of unreliable mappings between sounds and letters, but varying approaches to navigate said reliability.

> In English, there is a difference between the speling and the pronouncing of words, which make it importing to learn the the speling of a new word when hearing it . (AR5247103)
>
> It can give an indication of which sound you are lisening for in English - though not always (EN5503474)

There are also two examples where L1 English-speakers suggest that seeing English spelling can help clarify what you are hearing, but only on occasion. The rest of the responses are made in general reference to language learning. Additional insights into why participants do or do not value early exposure to written input are investigated within the specific categories described in section 5.5.2. The categories are presented in order of the number of L1 Arabic units and are contextualised in comparison to L1 English responses.

Clarifying and consolidating new language

The category with the largest number of response units is *clarifying and consolidating new language*, with 47 L1 Arabic and 54 L1 English units. All examples contain ideas around written language giving assurance of the correct form and increasing familiarity with the new word, as well as anchoring it better in memory. For example, the L1 Arabic responses include phrases such as: "ترسيخه" [*anchoring/grounding*] (AR5011969) and "هو بمثابة تثبيت للكلمة" *[It serves to fix/cement/root the word*] (AR5011969), which have connotations of establishing the word, in order not to forget. Another related example specifies this is a mental process, stating "حتى ترتكز في المخ" [*in order to anchor in the brain*] (AR5338811). Other responses state that OI was important "to know more" (AR5000812) and "to be familiar with that word" (AR5500050). Similarly, some of the L1 English-speakers report that OI was important "to build the concept of things in my brain" (EN5750994) and "in order to create a clearer mental representation in my mind" (EN5664502), as well as "spelling gives an anchor to schematize the sounds you hear" (EN5756084). All of which indicate awareness of psychological and linguistic concepts, and active integration of the written input to develop mental frameworks for new language.

Related ideas are reflected in the responses within the three subcategories, which specify that OI is important for *correct comprehension*, *correct pronunciation* and *decoding sounds and letters*. There are 15 L1 Arabic and 25 L1 English units that mention the importance of written input for *correct comprehension*. Across both language groups, some mention general improvement of understanding, as well as numerous examples where participants believe that they are more likely to mishear new words without seeing the spelling. Some specify that this is important in the context of discriminating similar sounding words and new sounds that are not part of your L1. Meanwhile, others express how written input helps identify and understand the meaning of the word, indicating that OI aids the parsing of both phonological and semantic information. Some then mention the importance of being able to identify a word when reading. Generally, there is a reluctance to rely on only one sensory input, particularly auditory senses, as this leads to uncertainty and room for error. In contrast, the written form offers certainty, confirmation, and "correctness".

للتأكد من صحة ما استمعت إليه (AR5000407)

[To be sure whether you heard correctly]

حتى أفرق بينها وبين كلمة أخرى شبيهة، وحتى أتعرف عليها في سياق القراءة (AR5339892)

[In order to differentiate between the word and other similar words, and so I recognise it in the context of reading]

لكي افهمها جيداً (AR5331395)

[To understand it well]

Without seeing it written, it can be hard to make sense of the sounds being made if the sounds are new to me. (EN5602966)

It prevents you learning it incorrectly incase you have mishead the pronounciation. (EN5499420)

it helps me to understand what the word means (EN5588274)

Rather than focusing on the reception of new language, 16 L1 Arabic and 11 L1 English units emphasise that written input is important for *correct production* of new words, both in speaking and writing. Responses across both language groups report frustration when feeling unsure about how to use a word correctly, and suggest that the written form is important for accurate and "proper" use. Overall, the L1 English responses focused more on pronunciation, whereas the L1 Arabic responses were more balanced in the importance they placed on both spoken and written usage. In fact, six mention only spelling or writing and not pronunciation.

because it will make it easier to understand how can i speak the word and how can i spell it. (AR5027570)

كي اعرف الاسم الصحيح والنطق (AR5622524)

[So I know the correct word and pronunciation]

Learning the spelling of a word allows you to use it in writing. I find it frustrating being unsure about the spelling of a word when I want to use it in writing. (EN5312548)

Moslty so you know, or can at least have a good guess, on how to pronounce it (EN5327303)

There are also some references to the standardisation of written language, in contrast to the variation found in spoken language, in the L1 English group. For example, the response below reveals ideologies around language variation, namely that only some people have accents, letters illustrate the sounds to be pronounced, and that deviation from a linguistic standard would be incorrect.

> it helps to register how i should pronounce the word correctly. If i always only listened to a person pronouncing it i may not say it right because the person speaking could have an accent and not clearly pronounce each sound that should be made when speaking clearly. Therefore without any spelling it would make me feel insecure about my speaking abilities and be worried my fluency would be affected by anothers mispronuncuation or natural talking habits. (EN5670727)

The final subcategory focuses on the importance of *decoding sounds from letters*. Within this subcategory, there are six L1 Arabic and 12 L1 English units. While there is some overlap with the other two subcategories, these responses focus more on sublexical analysis of the word. This is suggested to support memory, comprehension and production through matching sounds and letters, sounding letters out, and recognising patterns with words you already know. Several responses across both language groups mention the benefits of breaking multi-syllable words down into smaller sounds, through letters. They also report intentionally looking to understand the grapheme-phoneme correspondences in the new language through the spelling of words (EN5755300).

> حتى اربط بين املاء الكلمة وصوتها (AR5359752)
> *[In order to make a connection between the word's spelling and its sound]*
> i guess it's so that i can match sounds to letters (EN5615580)
> Because it helps for a learner to recognise patterns within the spelling system and how they interact with the sounds (EN5755300)

Some responses in the L1 English group mention the importance of learning "phonetically", which presumably relates to sounding out smaller units of words. However, this idea seems unclear, especially in reference to learning across different writing systems, as demonstrated in the quote below. This may relate to previously expressed ideas, where the importance of seeing the written input is *because* it is different from what you would expect based on the pronunciation. Therefore, as with an unfamiliar writing system, you would not intuitively be able to navigate the sound-symbol correspondences without exposure and instruction. Alternatively, this may relate to the general importance of decoding skills when acquiring literacy in a new script.

> Also helpful when the language being learnt has a different alphabet, so trying to learn phoentically is important. (EN5909763)

Remembering and memorising

The next largest category for the L1 Arabic group is broadly focused around the importance of OI for *remembering*. There are 17 L1 Arabic and 10 L1 English response units for this category, which all state that seeing the written form helps to remember new words and do not specify further details beyond that. These examples sometimes explicitly mention the belief that visual information is easier to retain.

> يساعدني جدا في تذكرها (AR5370467)
>
> *[It helps me a lot with remembering it]*
>
> If you have a visual indicator of the word (via its spelling), you can remember it more easily. (EN5998960)

An interesting distinction is then made when discussing the importance of OI for *memorising* in the L1 Arabic responses, as a separate concept to *remembering*. None of the L1 English responses refer to memorising new language, whereas there are 16 L1 Arabic responses that focus on this process in relation to written input. The creation of a separate category arose from discussion with my Egyptian collaborator, who related the importance placed on memorisation to the first step on *Bloom's Taxonomy* (Bloom, 1964), namely *recalling knowledge*, and the international influence of this taxonomy on teacher training.

While this taxonomy and its revised version (Krathwohl, 2002)[37] have had an undoubtedly far-reaching impact around the world, they have also been the source of much critique over the last 50 years (Seaman, 2011). In any case, it is of interest to note the potential permeation of ideas interpreted from this taxonomy in the L1 Arabic responses, where this foundational step of memorisation is supported by written input. This is likely to differ across cultural and educational contexts. For example, in the UK, rote memorisation may not be valued to the same extent (Watkins, 2000), which is potentially illustrated by the lack of reference to memorisation in the L1 English responses. Most examples highlighting memorisation simply state that witten input generally improves the process, or is important to memorise new words more easily or correctly. Some specify the opportunity to memorise both spoken and written forms simultaneously, potentially referencing ideas connected to dual-coding, previously mentioned in section 5.4.2.

To memorise it correctly (AR5003015)

احفظ النطق و الكتابة معا في نفس الوقت (AR5003023)

[I memorise the pronunciation and writing together at the same time]

Visual association and visualisation

The next largest category is *visual association*, where examples highlight the importance of seeing visual input to make a variety of different mental connections that support learning new words. There are 11 L1 Arabic and 12 L1 English units included in this category. Across both language groups, responses make reference to associations with the L1 or words that are already known to the participant. Others focus on having an additional visual association to accompany the sound of the word, sometimes referring to a memory advantage. There are also examples that include association to the meaning of the word, indicating that the combined orthographic and phonological information facilitates

[37] The revised version by Krathwohl (2002) changed the name of the first step to *remembering*, which is perhaps at odds with the distinction made within the present data. However, the collaborative interpretation of the data was that the terms did not appear to be used interchangeably, but rather denoted (subtly) different ideas and processes. Also, our frameworks of analysis are distinct.

connections with the semantic content of the new word. One of the distinctions between these responses and those in the category of *decoding letters and sounds* is that, here, participants tend to focus on the shape of the whole word rather than the individual letters.

حتى يكون هناك ربط بين الصوت ورسم الكلمة (AR5439128)

[So there is a connection between the sound and the shape of the word]

أحيانًا أربط بين الشكل والصوت أو المعنى (AR5359087)

[Sometimes I make a connection between the shape, the sound, and the meaning]

spellings provide a second sense to link the memory of the word to (sight as well as sound) (EN5592006)

You can make stronger connections between the word, its spelling and the concept it represents. (EN5998960)

Another category that emphasises visual processing and memory is that of *visualisation*. Rather than focusing on connecting the visual input with sounds and semantics, these responses refer to the importance of written input to hold an image of the word in the mind. There are 7 L1 Arabic and 8 L1 English response units in this category. The L1 Arabic responses tend not to use the word 'visualise' specifically, but instead refer to recalling, imagining, and drawing the shape of the word to remember it. Some responses mention this facilitates association, as well. In contrast, the L1 English responses straightforwardly state that written input helps to visualise or picture the word and at times connect this to memory and production.

حتى اتخيل شكل الكلمة (AR4986647)

[In order to imagine the shape of the word]

its help me to draw a vision on my mind and link between the word and the shape (AR5337282)

Helps me to see the word and visualise it (EN5672324)

you can picture the word spelling when you say it (EN6029761)

Speed of learning and learning styles

The next category could be seen to overlap with others that have been mentioned so far. However, a separate category was formed to represent the importance that some participants place on written input and the *rate* of learning of new words, as well as the manner. This category of *speed of learning* includes 6 L1 Arabic and no L1 English responses. All the L1 Arabic responses suggest that written input helps them to learn, memorise, or recognise new words quickly.

> لاحفظها بسرعة (AR5337115)
>
> *[To memorise it quickly]*
>
> That help me to learn quickly (AR5499774)

The final category to mention is *learning styles*, which includes responses where participants identify themselves as a visual or textual learner. This is presumably a reference to the widely used, but heavily criticised, Visual-Auditory-Kinaesthetic (VAK) framework (Dornyei & Ryan, 2015; Lethaby & Mayne, 2020; Papadatou-Pastou et al., 2021). There are three L1 Arabic and nine L1 English response units that are included in this category. Some mention this in contrast to being an auditory or aural learner; however, most simply report that they are a visual or textual learner (AR5059373). A couple of responses mention that their memory is therefore dependent on seeing something written down (EN5332693). There are also a couple of examples where participants state that written input is important for visual learners in general, but do not necessarily identify themselves as a visual learner (EN5909763). The usage of this terminology, especially without any further justification, demonstrates the prevalence of "learning styles" and the assumption that it is a widely-understood concept for teaching and learning.

> i am a visual learner (AR5059373)
>
> Visual learner, my memory seems quite dependent on having something in front of me to help with learning (EN5332693)
>
> It is helpful for people who are more visual learners. (EN5909763)

In summary, evidence has been presented exploring beliefs about why early exposure to written input when learning new words is important. Across both language groups most participants feel that it is generally important to see written input and that this can support the *clarifying and consolidating of new language*. Participants express the importance of seeing written input in order to understand and use new language correctly, as well as decode sounds and letters. Both groups also mention the importance of seeing written forms for *remembering* new words. This is more strongly represented in the L1 Arabic group, who also express ideas around the importance of OI for *memorising* and increasing *speed of learning*. In contrast, neither memorisation nor speed of learning was mentioned in the L1 English group.

Both groups highlight the importance of being exposed to visual input in the form of orthography. One category captures ideas around *visual association* and how it enhances connections in memory between the sounds, symbols, and semantics of a word, as well as associations with existing linguistic knowledge. A separate category focuses on being able to *visualise* and hold the shape of a word in the mind, which is also sometimes mentioned in relation to supporting memory. The focus on visual information is also apparent in the category of *learning styles*, where participants state being a visual or textual learner as the reason why seeing the word written down is important. Overall, these responses demonstrate varying beliefs that, at times, reflect ideas around orthographic depth, visual memory, dual-route approaches to reading, and theoretical models of education and learning. Additionally, some ideologies around linguistic standardisation and "correctness", in both receptive and productive language processing, have been revealed in connection to the perceived importance of exposure of written input.

5.5.2.2 Written input is not important

In contrast to the previous section, only nine (8%) L1 Arabic and 29 (24%) L1 English units express that early OI exposure is *not* important for learning new words in a language. None of the L1 Arabic responses make any specific references to either the English or Arabic script. Meanwhile, four L1 English responses mention spelling in English as an obstacle. One response states that seeing written input generally interferes with listening. Another response draws on a participant's experience with several different languages, including

English, resulting in the belief that spelling rarely reflects pronunciation. These and other negative or neutral views about the importance of written input are explored in categories where OI is not important because *it's too much to process*; *sound is more important/beneficial*, *spelling does not reflect pronunciation*, or based on *previous (language) learning experience*.

It's too much to process

The largest L1 Arabic category focuses on exposure to OI being *too much to process*. There are four L1 Arabic and seven L1 English units in this category. A couple of responses indicate that it is simply not necessary during early exposure, whereas others express that the additional written input can be distracting and result in only focusing on the spelling of the word. In both language groups participants indicate difficulty trying to learn, memorise, and process both spellings and sounds simultaneously, implying cognitive overload.

> ممكن أن أتشتت عن استماع الكلمة بالشكل الصحيح وأركز على شكها فقط (AR5212731)
>
> *[I can be distracted from hearing the word correctly and only focus on the shape]*
>
> لانه يصعب علي المهمة في حفظ الاملاء للاحرف وحفظ الكلمة المستمعة (AR5450539)
>
> *[Because it makes it hard for me to memorise the spelling of the letter and memorise the sound of the words]*
>
> written and hering the word is more to process (EN5759735)
>
> Easier to concentrate on one aspect of the word first. ie just hearing it or just seeing it (EN5914367)

Sound is more important/beneficial

The largest category for the L1 English groups suggests that *focusing on sound is more important/beneficial*, over orthographic input. There are two L1 Arabic and 10 L1 English units in this category[38]. The L1 Arabic responses broadly state that sound is more important. Meanwhile, additional detail from the L1 English group suggests that, particularly when thinking about first exposure and order of input presentation, it is better to listen first and then see the spelling later (EN5337460). Some responses connect focusing on the sounds of

[38] The remaining categories only have one or two examples from L1 Arabic-speakers, reflecting the limited expression of negative beliefs around the importance of early exposure to written input. Therefore, the descriptions and examples for the next sections are based on the L1 English data.

the word with improved pronunciation and recognition. Another participant mentions paying attention to the details of the sounds in terms of syllables and specific vowels.

> sound more important (AR5365353)
> Maybe the first time it's important to try to focus only on the sounds, but by the second time, I do prefer to have the written version too. (EN5337460)
> If the focus is purely on sound, hearing and then repeating orally is more beneficial in consistently replicating and identifying the sound/word (EN5434242)

An alternative perspective is presented by a participant in the L1 English group, who acknowledges that learning based on listening, without OI, does not make it easier to learn a new word. Rather, the word would be better established in memory because more effort is required to initially encode the information. This may align with ideas related to *desirable difficulty*, in both psychology and language learning research (E. L. Bjork & Bjork, 2011; R. A. Bjork & Bjork, 2020; Suzuki et al., 2019). While there are many forms of undesirable difficulty, these authors note a particular level of difficulty, relative to prior learning and background knowledge, which strengthens the encoding and retrieval processes that underpin learning, comprehension, and use.

> When learning a new word, just hearing it rather than hearing it and seeing it requires more effort to retain therefore I think you'd be able to remember it more easily as more effort has gone into remembering it. (EN5706491)

Spelling does not reflect pronunciation

The next category to mention focuses on the belief that *spelling does not reflect pronunciation*. There are two L1 Arabic and eight L1 English units in this category. Participants refer to spelling as misleading and confusing, stating that it is not always a good guide to pronunciation. Three of the L1 English responses refer to English spelling specifically, demonstrating how this belief is likely related to L1 literacy experience (EN5659326). Some examples also refer to the orthographic depth of different writing systems and the processing demands associated with learning incongruent spellings (AR5336140).

> beacuse sometimes writing does not refelct how we speak the language. there may be little correspondnce between the writing systme and the sound of a language. (AR5336140)
>
> In english the spelling doesn't always match up with pronunciation so it could make it difficult. (EN5659326)
>
> Spelling doesn't always help with pronunciation, sometimes it's easier to listen and repeat without thinking about how to spell it (EN5750023)

Previous (language) learning experience

The final category includes responses that focus on *previous experiences of language learning*. There is one L1 Arabic response in this category and four L1 English responses. The Arabic response is another reference to learning styles, where they identify themselves as an auditory learner "انا شخصية سمعية" (AR5000759) *[I am an auditory person]*. It is included in this category for clarity between the two types of affirmative or negative responses, as well as the assumption that this belief is based on previous learning experiences. The L1 English responses refer to broad beliefs about the order of acquisition, such as children learning languages primarily through listening at the start. Other responses express personal preferences or observations about what has worked when learning languages in the past. The first example below (EN5344305) also indicates that oral ability is not dependent on literacy ability, which contrasts some of the beliefs articulated in 5.5.2.1, particularly the importance of written input to produce spoken language. However, this participant also echoes the importance of OI exposure for producing (correct) written language.

> Because, that's how I learnt French. However, my spelling is awful so it probably is important to do so but it wasn't something that stopped me learning to speak. (EN5344305)
>
> Because when you first learn a language, as a young child, its mostly done through listening (EN5659195)

In summary, less than 10% of L1 Arabic-speakers and almost 25% of L1 English-speakers believe that it is not important to see written forms early on, when learning words in a new language. Across both language groups, participants link early exposure to multimodal input to cognitive overload, indicating that the visual information is given priority and distracts from the auditory input. This perspective complements beliefs that sound is broadly more important or beneficial during the early stages of language learning. Participants also express beliefs about unreliable sound-spelling correspondences, with some mention of additional processing demands. This is notably more present in L1 English responses, as is reference to previous experiences of language learning.

5.5.3 Summary of perceived OI importance

Responses to Q7, of whether written input was important when learning a language and hearing a word for the first time, revealed that most participants in both language groups believed that it was, indeed, important. This tendency was particularly strong in the L1 Arabic group, where only 12% of all responses indicated that OI was not important, compared to 28% of L1 English responses. Similar proportions were also reflected in the open responses to Q8, where less than 10% of L1 Arabic responses expressed that written input was not important, in contrast to 25% of L1 English responses. The fact that there were almost double the number of response units included in the qualitative analysis for the L1 Arabic group for Q8 compared to Q6 demonstrates that participants have stronger or clearer views in relation to this question. This underscores the importance placed on written input and that participants are aware of specific ideas and reasons to justify such a belief.

Further light was shed on these group tendencies through exploration of the open responses to Q8 and the inductive formation of categories based on the data. Figures 5.10 and 5.11 display the proportion of response units that were included within each identified category and subcategory, by language background. These treemaps illustrate the frequencies described above, showing the larger proportion of affirmative responses in blue and smaller proportion of negative responses in orange. While proportions and frequencies differ across the two groups, there is commonality in terms of the categories mentioned and the emphasis on written input being important for *clarifying and consolidating new language*. Participants across both groups articulate the importance of written input for understanding and using new words correctly, as well as being able to break down and match up individual sounds and letters. This is contrasted with the smaller number of negative responses holding the opposing view, where written forms are unreliable and not a good guide for pronunciation, as well as potentially distracting from, rather than helping attend to, the sounds of a new word.

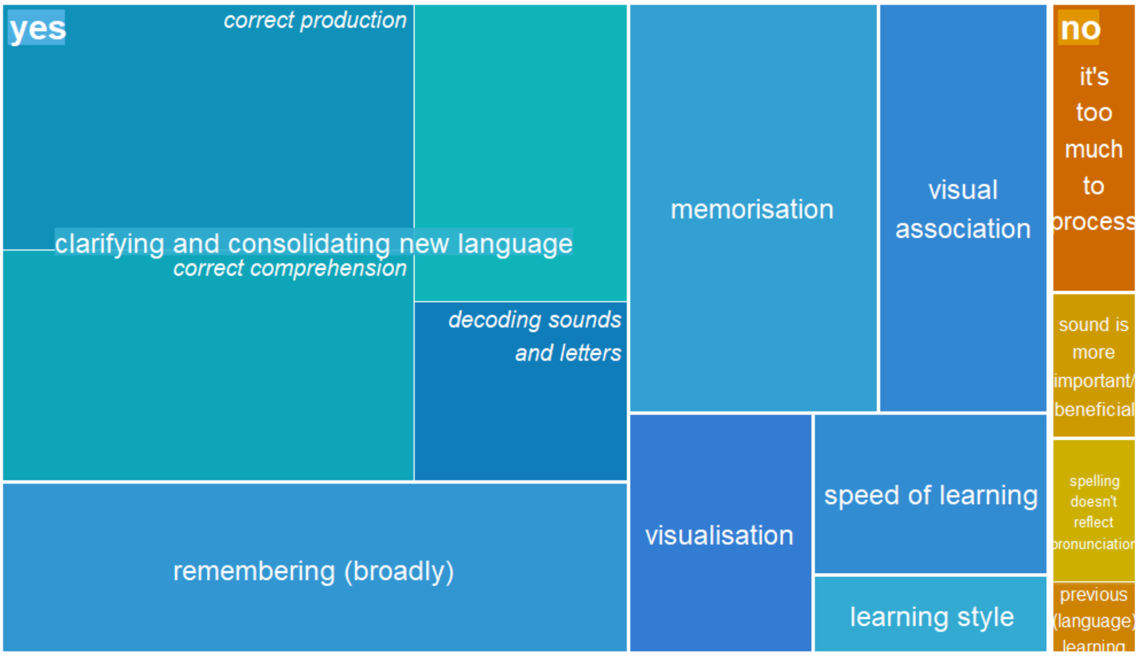


Figure 5.12: Treemap of L1 Arabic beliefs about why early exposure to OI is important

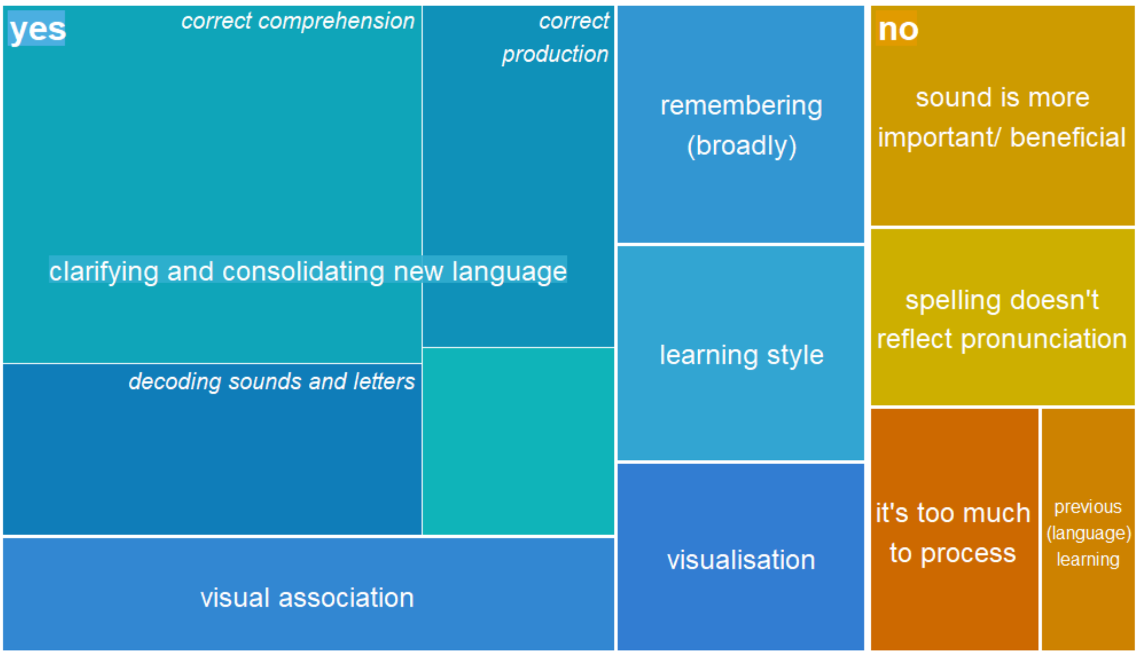


Figure 5.13: Treemap of L1 English beliefs about why early exposure to OI is important

Additional positive beliefs around the importance of OI included general support when *remembering* words. This was found across both language groups, but more so in the L1 Arabic group. A related category was also formed on the basis of the L1 Arabic data alone, to capture the frequent references to *memorisation,* specifically. As mentioned, when discussing these responses with an Arabic-speaking collaborator, it was hypothesised the stronger emphasis on both remembering and memorisation may reflect cultural differences in educational trends, such as rote learning techniques. For example, these responses were linked to the first step on *Bloom's Taxonomy*, where the foundational level of learning and knowledge is reflected in the ability to memorise and recall information. Some L1 Arabic-speakers also focused on *speed of learning*. The importance of written input to rapidly learn vocabulary indicates a particular learning goal, which according to Tyler (2019) may have a limiting influence on L2 phonological development. However, this is speculative. Again, it is worth noting that the lack of explicit reference does not mean that these processes or beliefs were not held by L1 English-speakers.

Participants in both groups also refer to the importance of visual input for creating *visual associations* and holding an image of the word in the mind through *visualisation*. These examples often make a connection with the storage of new words in memory. The associations mentioned vary from connecting the sound or meaning of the word with the image of the spelling, or forming associations with existing linguistic knowledge. While most participants who mention multimodal input in this way suggest it is advantageous when learning new words, there is a minority of participants who believe that this same information is *too much to process* all at once. Perhaps, this reflects individual differences, or that those who rely heavily on the written forms are not aware of the auditory details that are missed through split attention, as they feel more confident in their ability to recall what they have comprehended.

The potential issues of forging associations between new and existing linguistic knowledge, based on familiar sound-symbol correspondences, have been illustrated by research into orthographic influence on L2 phonology, particularly when there is inconsistency both within and between languages (Bassetti & Atkinson, 2015; Erdener & Burnham, 2005; Hayes-Harb & Barrios, 2021). Awareness of this effect is evidenced in the negative category which highlights that *spelling does not reflect pronunciation*. Then, returning to the point of individual differences and styles, it is of note that some participants, across both language groups, identify themselves as visual/textual learners or auditory learners. A critique of the application of learning styles like this is that it fails to communicate the range of perceptual senses we all use to learn, even if there is a general preference for a more visual, auditory, or kinaesthetic approach. Additionally, when learners identify themselves in this way, it can lead to situations where someone is closed to the idea of learning through senses other than their identified perceptual style. Thus, such beliefs may have consequences for how "visual/textual" learners attend to auditory input.

In relation to the research question stated at the start of this chapter, these results demonstrate that the majority of participants believe written input is influential for their learning. Most participants, across both language groups, believe that early exposure to written input is important when learning new words in a language, emphasising the benefits for clarifying and consolidating new language in relation to comprehension, memory and production of new words, amongst other things. However, in both language groups there is also a minority of participants who do not hold this view, based on beliefs about processing demands, unreliable sound-spelling correspondences and past learning experiences. The final section of this chapter discusses these beliefs, together with the perceived influence of OI during the task, to provide an overview of perceptions of orthographic influence.

5.6 Concluding remarks

Frequencies of participant responses to all closed questions, related to the experiment specifically and language learning more generally, clearly demonstrated that most participants in both language groups perceived written input to facilitate the learning of new words. In the context of the experimental task, there was a preference for written input to be presented in English, and the majority of L1 Arabic-speakers (61%) reported more difficulty learning without any written input. This provides the first indication that participants were not aware of an inhibitory effect of OI, particularly when learning words containing a confusable contrast. These frequencies also suggested that participants viewed written input in English or Arabic script to have a different impact on learning, contrasting the results of chapter 4. The focus on the positive influence of English OI further demonstrated that L1

Arabic participants were familiar with the English script, and perhaps accustomed to learning English words in this script. Meanwhile, the mixed responses in relation to Arabic OI demonstrated more uncertainty about the influence of their L1 script. Most L1 English participants indicated that the unfamiliarity of Arabic spelling made no difference to word learning, and only a small number perceived it to be more difficult. Some participants also reported not seeing any written input at all during the experiment. Most likely, this was only a small number of participants, especially when considered in light of participant open responses. However, this would be worth exploring more systematically in future research.

The qualitative content analysis of participant open responses also revealed an overwhelmingly positive perception of written input, in both the experiment and language learning more broadly. This was particularly prominent in the responses of the L1 Arabic group, where over 90% indicated that exposure to written input helped them in the experiment and was important for learning new words more broadly. Meanwhile, in the L1 English group almost 25% of responses expressed that written input had a negative impact or made no difference to word learning. With regards to the specific content of the responses, it is important to acknowledge that understanding of participants' perceptions may be lost in translation. Firstly, I was navigating translations as an L2 Arabic-speaker, making sense of short responses, from a range of Arabic dialects and cultural contexts. Secondly, around half of all responses from the L1 Arabic group were given in English, thus these participants were also drawing on multiple linguistic repertoires, with varying degrees of proficiency, to communicate their ideas. In order to address this, all Arabic responses were discussed with an Egyptian bilingual colleague, who had experience researching language and education with L1 Arabic-speakers, as well as QCA. Additionally, a satisfactory level of intercoder agreement supports the categories reported here.

Looking at the results from both Q6 and Q8, regarding perceptions about the experiment and language learning more broadly, there were consistent references to the positive influence of written input in relation to remembering and being sure about word forms. While there were some references to the difficulty of /f-v/ contrast words and the support that OI can offer to differentiate similar sounds in the L1 Arabic group, this was only mentioned in a minority of responses. Most participants expressed a variety of perceptions, strategies, and beliefs around the relationship between letters, sounds and mental processes. The quality and quantity of the responses to Q8 were improved compared to Q6, indicating that participants had more difficulty understanding the question and articulating their reflections on written input during the experiment. However, in response to Q6, similar ideas were expressed across both groups, including ideas about the relationship between sounds and letters in each script, and the influence of multimodal input on memory and processing. Reflections on learning during the study appeared to relate to broader beliefs and goals, in relation to language learning. For example, some stated that the lack of systematicity in English spelling was a reason to be exposed to it early, whereas others suggested this was something to avoid. The importance placed on OI may reflect that some participants prioritised being able to accurately use a word in writing as quickly as possible.

The idea of written language being more correct, concrete, and certain was repeated throughout many articulations, revealing beliefs about the nature of writing and how it represents language, as well as a mistrust of learning through hearing alone. Additionally, a strong connection was made between visual input and improved memory of items. The focus on remembering, memorising and rapidly learning new words was more apparent in the L1 Arabic responses compared to the L1 English group. This was hypothesised to be related to different pedagogical approaches in cultural contexts, where a different value may be placed on rote learning and recalling new knowledge as a foundational educational step. Ideas around correctness and importance of exposure to written forms in language instruction could also reflect cultural values around sacred texts and the complexities of diglossic literacy development, which have been explored in Arabic linguistics, sociolinguistics and anthropology (Haeri, 2000; Saiegh-Haddad & Spolsky, 2014), as noted in chapter 2.3.1. For example, most Arabic speakers are taught literacy in Classical Arabic and Modern Standard Arabic through religious and state education, where high prestige is attached to these varieties. In contrast, literacy in everyday spoken varieties is typically non-standardised and informal, leading to a different understanding of the relationship between spoken and written linguistic expression than in European languages. Furthermore, when communicating through computer-mediated social platforms, speakers of Arabic dialects have had to navigate technological limitations and devices which only offer the roman alphabet to create uncodified transliterations, leading to the evolution of phenomena such as *Arabizi* (Allehaiby, 2013). However, as technology advances and the internet is better able to support different languages and writing systems, this is changing.

Additionally, studies into Arabic lexical processing from a psycholinguistic perspective have drawn attention to the ways that spoken and written word recognition differ in comparison to English. One such study by Boudelaa (2014) uses priming and neuro-imaging evidence to demonstrate that Arabic lexical representations are organised in terms of roots and patterns. This may relate to references in the L1 Arabic responses to <شكل الكلمة> *[the word shape],* indicating that participants were looking to encode patterns and the shape of the whole word, perhaps more so than focusing on the first letter, which was evident in L1 English but not Arabic responses. The transfer of L1 literacy processes and reading strategies to L2 learning, specifically in the context of L1 Arabic learners of L2 English, has also been demonstrated in the context of the different treatment of vowels in Arabic and English orthography (Hayes-Harb, 2006; Saigh & Schmitt, 2012). This distinction was mentioned in a quote at the start of section 5.4.2.1, demonstrating an awareness of the levels of information encoded across orthographies and the need to potentially navigate different reading strategies. Further to this, the increased attention to the inconsistencies of English spelling and the benefits of focusing on sounds, reflected in L1 English responses, also likely reflects L1 literacy experience.

Across both language groups, there was an overriding idea that written input supported the clarification and consolidation of new language in memory. This draws attention to the ways participants perceive their processes around the encoding of novel lexical items, primarily based on the visual information they are presented with, where only a minority noted the potential cost for attention to auditory input. The low awareness of insufficient attention to acoustic cues may offer explanatory insight into the inhibitory effects of OI when lexically encoding the distinction between confusable contrasts, reported in chapter 4. Participant responses emphasised using written forms to help clarify similar sounding words, visualise the word in order to recall it later, and create visually-based associations. These approaches to the task may have distracted from necessary phonetic detail to encode precise phonological distinctions, impacting the accuracy of spoken language processing later. Perhaps, participants are better able to recall the words, or have a sense of stronger mental representation, when they have seen the written form. However, they are not in a position to assess the accuracy of the representation they have encoded.

This detailed account of both the analytical procedure and interpretation of responses is offered in advocacy of increased mixed method and qualitative approaches in the field of orthographic influence on L2 phonology. The reflections from participants presented here demonstrate an awareness of orthographic influence on L2 acquisition, as well as active direction and suppression of attention and the priority given to rapidly integrating the visual information from orthography in lexical representations. These findings also suggest that the way that participants interact with the input they are presented with is related to broader beliefs about the relationship between sounds and symbols, ideologies of "correctness", and educational experiences. These preliminary investigations into the relationship, or contrast, between participant perceptions and performance in psycholinguistic studies demonstrates great scope for further research. The next chapter takes a closer look at the strategies participants employed in the experimental task, particularly in relation to phonology and orthography, and then investigates strategy use in relation to their performance in the matching task.

Chapter 6: Participant strategies and written input

6.1 Introduction

This chapter presents the analysis of reported language learning strategies (LLS) during the word learning and matching tasks, based on Likert-scale responses to a LLS inventory and open responses to the post-test questionnaire. Whilst open responses referred to strategies broadly, the inventory is designed to focus on strategies that relate to phonology and orthography. This is with the aim of understanding (a) if participants make use of overtly phonological and orthographic strategies, (b) if participants make use of the accompanying written input, (c) if strategy use predicts higher accuracy lexically encoding the target contrasts, and (d) if strategy use is related to participant characteristics, such as L2 proficiency. Additionally, this chapter responds to calls for more domain-specific LLS, such as pronunciation and phonology (Oxford, 2016; Pawlak, 2021), by presenting a novel LLS inventory, focused on phonology and orthography. As with the previous chapter, the data from the two language groups are analysed separately and then discussed comparatively, with a focus on the L1 Arabic-speakers. The research question being addressed is:

What is the relationship between learners' language learning strategies and lexical encoding of L2 phonological contrasts in novel words?

Initial analysis of participants' accuracy on the matching task, reported in chapter 4, revealed that performance was significantly worse when words differing by the confusable /f-v/ contrast were taught with either English or Arabic OI, in comparison to audio-only exposure. L2 English proficiency significantly predicted the accuracy of participants' responses, where higher proficiency learners performed better than those with lower English proficiency. However, even the highest proficiency learners struggled to lexically encode the /f-v/ contrast when accompanied by written input. Chapter 5 went on to demonstrate that participants' perceptions and beliefs around the importance and benefits of written input may be at odds with performance in the matching task. Additionally, it was evident that participants were consciously employing varied LLS, motivated by distinct learning beliefs and goals. To further investigate the discrepancy between participant perceptions around orthographic influence and ability to lexically encode a confusable contrast with written input, this chapter explores the LLS that participants employed during the study in more depth. This is partly to understand the ways in which participants engaged with the experiment, as well as assess the quality of the data and conclusions arrived at thus far. However, the primary aim is to identify participant strategies and explore the relationship between strategy use and task performance.

This chapter offers a novel perspective within the field of orthographic influence on phonological acquisition, as well as LLS research. While research into LLS has been established for around 30 years, with the 1980s and early 1990s being particularly productive periods, there has been limited research into L2 phonology and pronunciation (Chamot, 2019; Oxford, 2016; Pawlak & Szyszka, 2018). Additionally, I am only aware of one study, by Bassetti and colleagues (2020), in the L2 orthography and phonology literature that incorporates LLS assessment into their study. As part of their investigation into the effects of individual differences (ID) and orthographic influence on L2 production and phonological awareness, participants completed an adapted version of the *Pronunciation Learning Strategy Inventory* (Berkil, 2008), to obtain the number of strategies used. However, they did not find this to be a significant predictor of participants' consonant or vowel duration in speech production. The lack of statistical significance across the speaker-level variables in the study was highlighted to demonstrate the need, not only for more research into ID variables in relation to orthographic influence, but also for alternative measures of such variables. This chapter focuses on the mixed methods analysis of participant responses to a novel strategy inventory and open questionnaire responses about strategies used, specifically in reference to the word learning and matching tasks in the study.

6.2 Research context

As mentioned, this chapter heeds calls for more empirical investigations into the strategies used for specific sub-skills and how they are deployed in the context of learning tasks (Oxford, 2016; Pawlak, 2021). Oxford (2016) notes the paucity of research into strategies for L2 phonology, as well as limited in-depth research into pronunciation and listening strategies. Pawlak and Szyszka (2018) have also pointed to the scarcity of research into pronunciation learning strategies (PLS), which they do not distinguish from phonological strategies. In their review of the literature, they note that findings are often contradictory and difficult to draw clear conclusions from. This is partly due to a lack of consistency in the measurement of PLS and a reliance on correlational analysis, which limits the ability to make claims about cause-and-effect relationships.

To illustrate the point, a few key studies that are drawn on for the present study are outlined in table 6.1 (see Pawlak & Szyszka (2018) for a more exhaustive review). Almost all these studies use a different inventory to assess pronunciation strategies with varying outcomes, both in terms of identifying strategies and exploring the relationship between strategy usage and pronunciation ability. A point of unity is that almost all take inspiration from the hugely popular *Strategy Inventory for Language Learning* (SILL) (Oxford, 1990). Originally a self-evaluation tool for L2 learners, which divided strategies into six subscales (memory, cognitive, compensation, metacognitive, affective, and social), it has since attracted both international acclaim and criticism. In the SILL, learners respond to items, such as "I try to find patterns in English", with a five-point Likert-scale, ranging from: *never or almost never true of me, generally not true of me, somewhat true of me, generally true of me,* and *always or almost always true of me*. Oxford (1996, p. 31) outlines the subscales as:

1. *Memory strategies*, such as grouping, imagery, rhyming, and structured reviewing (9 items).
2. *Cognitive strategies*, such as reasoning, analysing, summarising (all reflective of deep processing), as well as general practising (14 items).
3. *Compensation strategies* (to compensate for limited knowledge), such as guessing meanings from the context in reading and listening and using synonyms and gestures to convey meaning when the precise expression is not known (6 items).
4. *Metacognitive strategies*, such as paying attention, consciously searching for practice opportunities, planning for language tasks, self-evaluating one's progress, and monitoring errors (9 items).
5. *Affective (emotional, motivation-related) strategies*, such as anxiety reduction, self-encouragement, and self-reward (6 items).

6. *Social strategies*, such as asking questions, cooperating with native speakers of the language, and becoming culturally aware (6 items).

The SILL has faced opposition due to the limitations of self-report Likert-scales to measure strategy frequency, as well as statements being outdated or inappropriate in different sociocultural contexts. However, as argued by Amerstorfer (2018), SILL still proves to be a useful tool for learners, teachers, and researchers, alike. In particular, she draws attention to the flexibility of SILL to be integrated into mixed methods research and adapted for specific contexts. The studies in Table 6.1 demonstrate how the SILL has been adapted to focus on domain-specific strategies, such as pronunciation. Another commonality between these key studies is the recruitment of university students, and often Turkish and Polish learners of English studying linguistics or philology. It is worth noting that SILL has also influenced research looking at literacy learning across writing systems, namely strategies used by L1 English learners of L2 Chinese characters. Research using the *Character Learning Strategy Inventory* (Shen, 2005; Sung & Wu, 2011) was not drawn on in the design of the present study, as the research did not focus on phonological acquisition. However, their findings regarding orthographic-knowledge-based cognitive strategies are discussed alongside the quantitative results presented in section 6.5.

Table 6.1: Sample of key studies investigating pronunciation learning strategies

Author(s)	Participants	Study	Key findings
Peterson (1997)	11 L1 English speakers learning L2 Spanish (beginner to advanced)	Diaries and interviews focused on pronunciation learning strategies (PLS).	12 pronunciation strategies and 43 specific tactics, categorised based on Oxford (1990). Positive relationship between strategy use and pronunciation ability.
Berkil (2008)	40 L1 Turkish learners of L3 English (upper-interme diate-advanced)	Explorative design with *Strategy Inventory for Learning Pronunciation* (SILP) adapted from Peterson (1997) and two oral elicitation tasks.	11 new pronunciation learning strategies were added to Peterson's inventory. No difference in strategy use in relation to proficiency.
Pawlak (2010)	80 L1 Polish	Development of the	Validated instrument with

	students of L2 English (advanced)	*Pronunciation Learning Strategy Survey* (PLSS) (Cronbach $\alpha = 0.69$).	60 likert-scale statements and open-ended items. Statements adapted from SILL (Oxford, 1990).
Całka (2011)	74 L1 Polish students of L2 English (varied proficiency)	PLS questionnaire based on Oxford (1990) and Peterson (1997).	Prevalence of cognitive strategies and metacognitive strategies.
Rokoszewsk (2012)	63 L1 Polish students of L2 English (varied proficiency)	PLS questionnaire from Całka (2011) with an open question (Cronbach $\alpha = 0.89$). Perception and production test of English vowels.	PLS used occasionally, mostly cognitive and metacognitive strategies. No relationship between PLS use and perception but positive relationship with production.

In a similar vein, the present study draws on the SILL (Oxford, 1990) for relevant strategies, including listening, reading, writing, vocabulary learning and pronunciation strategies, as well as previous questionnaires developed to specifically assess PLS (Berkil, 2008; Rokoszewska, 2012). Like Pawlak and Szyszka (2018), these studies generally do not distinguish between pronunciation and phonology strategies, such as the inclusion of strategies like *representing sounds in memory* and *analysing the sound system*. Therefore, following the example of Bassetti et al. (2020), these questionnaires were chosen to direct the formation of a strategy inventory that focused on both phonology and orthography. It is noteworthy that the PLS inventories mentioned in table 6.1 often include items related to orthographic influence, such as reading aloud or visualising a transcription of the word. However, there is no discussion of the relationship between phonology and orthography, as well as conflation of strategies involving sounds and letters. For example, Peterson's (1997) modified SILL includes the item "when first learning another language, I decide to learn the sounds or the alphabet right away." This demonstrates the need for an inventory that considers the relationship between orthography and phonology more systematically.

While there is limited empirical and theoretical basis when it comes to identifying strategies relevant to L2 phonology, orthography, and their use in language learning, the following preliminary predictions are made:

1. Participants will report using strategies related to both written and audio input for the purpose of L2 phonology and word learning.
2. Cognitive and metacognitive strategies will be reported in the context of the language learning task in this study.
3. Strategy usage will predict matching task accuracy and relate to L2 proficiency.

Overall, the present study extends existing research to explore the influence of orthography on phonological acquisition, word learning and strategy use, as well as the strategies used by a different population; namely, L1 Arabic learners of L2 English. It also presents a novel LLS inventory and a concurrent design, with qualitative insights into learners' reflections related to LLS.

6.3 Methodology

An overview of the methodology and research instruments relevant to this chapter are reiterated for clarity. The development of the *Phonology and Orthography Language Learning Strategies* (POLLS) inventory is presented in chapter 3.6.8, including the formation of strategy statements and comparison between the POLLS inventory items and existing LLS inventories (Berkil, 2008; Oxford, 1990, 2016; Rokoszewska, 2012). While strategies from existing inventories were relevant, they often referred to language learning in classroom and naturalistic settings, rather than an experimental context. Therefore, the starting point of the POLLS inventory was analysis of questionnaire responses from my master's dissertation and interview data from piloting. This decision was based on the high relevance of the strategies mentioned, with specific reference to simultaneous written and spoken input when lexically encoding a confusable phonological contrast during a word learning experiment.

It was important to keep the inventory comprehensive but brief, so as to avoid attrition in participation. Therefore, the final 23 statements by no means fully encompass all strategies used in the task, but focus on the topic of this research project, namely the influence of written input on L2 phonology. Table 6.2 provides the wording for the open question and strategy statements, in the order they occurred during the post-test questionnaire and with corresponding question number. The open question was placed before the strategy statements, to avoid influencing participants' initial reflections on the task. The lack of distinction between memory and cognitive strategies, found in previous inventories (Berkil, 2008; Oxford, 1990; Rokoszewska, 2012), aligns with developments in the field, such as the inclusion of memory strategies within cognitive strategies in the Oxford's S2R model (2011, 2016). Additionally, the imbalance between cognitive, metacognitive strategy statements reflects the content of the master's questionnaire and pilot data. This echoes previous studies, which report high popularity of cognitive PLS, in comparison to other domains (Całka, 2011; Pawlak & Szyszka, 2018; Rokoszewska, 2012).

Table 6.2: Questions from the post-test questionnaire related to strategies

Domain	Question	Type
General	13) How did you try to learn or remember the words?	Open
Cognitive	14) I created association with words or things I already know. 15) I put the word in a context to remember it (e.g., a sentence, a story, a rhyme). 16) I made a mental image or imagined additional connections to help me remember. 17) I visualised the spelling of the words in Arabic in my mind. 18) I visualised the spelling of the word in English in my mind. 19) I repeated the words out loud or in my head. 20) I thought about or practised mouth positions to pronounce the words. 21) I thought about an action or movement to help remember the words. 21) I broke words down into syllables or sounds. 22) I tried to find patterns in the new words or sounds. 23) I tried to connect sounds and letters. 24) I tried to remember the first or last sounds. 25) I looked for similarities and contrasts between the pronunciation of the words and the languages I know. 26) I grouped similar sounding words. 27) I used the English spelling to distinguish between similar sounds. 28) I used the Arabic spelling to distinguish between similar sounds.	Likert scale *[always, almost always, sometimes, almost never, never]*

Metacognitive	29) I purposefully ignored the English spelling. 30) I purposefully ignored the Arabic spelling. 31) I found ways to test my memory and recall the new words. 32) I thought about my progress in learning the new words. 33) I noticed my mistakes and used that information to help me do better.
Affective	34) I noticed I felt more relaxed or confident when I saw the Arabic spelling. 35) I noticed I felt more relaxed or confident when I saw the English spelling.

In line with the instrument implemented by Rokoszewska (2012), the questionnaire included an open question, and the Likert-scale options that accompanied each statement were *always*, *almost always*, *sometimes*, *almost never*, and *never*. The reliability of this instrument was measured using Cronbach's alpha (α= 0.75), indicating an acceptable level of reliability (Dörnyei & Dewaele, 2022). To ensure that no important strategies were missed in the inventory, and to gain additional context for participant responses, Q13 was also included as an open question. As in chapter 5, all items were provided in the first language of participants. However, around half of participants in the L1 Arabic group chose to respond in English to the open question. A sample of responses from both language groups for Q13 are provided in table 6.3, to offer some insight into the type of data elicited. The same exclusion criteria were applied to the questionnaire data as in previous analyses, leaving the data from 114 L1 Arabic and 117 L1 English participants.

Table 6.3: Random sample of responses to Q13 by both L1 Arabic and English-speakers

Q 13: How did you try to learn or remember the words? **كيف حاولت تعلم أو تذكر الكلمات ؟**		
Group	**Participant ID**	**Response**
Arabic	5000814	by listening and repeating
Arabic	5003023	بربطها باشياء اخرى في الحياة اليومية مثلا كلمة fumel تبدا بحرف الفاء و تشبه الوردة "flower"

		[*By linking with other things in everyday life, for example the word "fumel" began with the letter "F" and looks like a flower "flower"*]
Arabic	5027570	i tried to think about how that person say it and guessed the spelling upon my experience
Arabic	5037277	I was trying to remember the spelling.. I mean the sound 🔊
Arabic	5038591	ربطتها بكلمات عربية [*I linked them to Arabic words*]
English	5308502	made up stories about them
English	5312160	By sound associations with word I already know.
English	5592006	said them out load to myself while looking at the picture, and repeated until image disappeared
English	5384030	look at the shape and see if there was any link to the sound. ie the image for makum looked like an m shape
English	5663887	analogies of things they reminded me of. e.g. the famel made me think of camel humps, so in my mind it was a Fat cAMEL

The following sections analyse this data in order to establish (1) what language learning strategies participants report using and (2) what the relationship is between language learning strategies and lexical encoding of L2 phonological contrasts. Both qualitative and quantitative analyses are applied to identify what strategies were used during the study. This takes the form of a Qualitative Content Analysis (QCA), followed by descriptive frequencies of responses to the POLLS inventory. A Principal Components Analysis (PCA) is then conducted to reduce the dimensions of the data, offering another perspective into the key components underlying the reported strategies. Finally, the extracted strategy components are analysed using correlational and linear regression analyses, to explore the relationship with lexical encoding accuracy and influential variables that have been highlighted in previous chapters, such as proficiency.

6.4. Qualitative Analysis

6.4.1 Analytical approach

It is common to approach open items on this kind of questionnaire with a form of content analysis (Dörnyei & Dewaele, 2022). Therefore, the same approach was taken to qualitatively analyse open responses as in chapter 5.3, namely QCA following the procedure laid out by Mayring (2014). The research question and theoretical background outlined in this chapter indicate the material relevant for analysis should refer to strategies used in relation to learning the phonological form of a word and modality of input during early word learning. As such, the direction of analysis is to arrive at statements on participants' strategy usage when learning new word forms, particularly in relation to phonology and orthography.

As before, existing research and relevant theoretical considerations are used to formulate selection criteria, category definitions and level of abstraction before analysis. Table 6.4 shows how these are defined for the open responses to Q13. For example, research into LLS has provided definitions, outlined in chapter 2.1.3.3, that are drawn upon for this analysis (Griffiths, 2018; Oxford, 2016). Additionally, research related to PLS highlights the particular use of cognitive and metacognitive strategies, thus it is anticipated that these domains will be reflected here, as well (Całka, 2011; Rokoszewska, 2012). There is also consideration for research related to L2 vocabulary, reading, and listening strategies (Oxford, 2016), based on the intersecting strategies related to the influence of orthography. However, due to the artificiality of both the learning context and the words themselves, the strategies relevant to this study focus on learning the decontextualised form of the word, rather than situated in an authentic context or based on usage.

Table 6.4: Overview of selection criteria, category definitions and level of abstraction for analysing responses to Q13

	Q13
Coding unit	Clear semantic units within participant open responses.
Context unit	Whole response to relevant question, the wording of the question, responses in the background and debrief questionnaire broadly.
Recording unit	Responses for 117 L1 English and 114 L1 Arabic participants

Category definition	• Actions chosen by learners and purposefully applied to the experimental task with the goal of learning. • Strategies can be flexibly combined in clusters or chains. • Strategies that refer to the modality of input. • Strategies chosen with the goal of learning the (phonological) form of new words.
Level of abstraction	Specific strategies used by learners, understood within general domains of cognitive, motivational, affective and social, and associated metastrategies associated with SR2 model (Oxford, 2016). Strategies are restricted to those relevant to learning the form of the word, as usage does not apply to the experimental task.

Following the same procedure as reported in chapter 5.3, responses remained in their original typed form and were not adjusted or translated before coding. The codes were all in English, for continuity across the whole dataset. After coding the first ~50% of L1 English responses, 13 categories were formed, which were improved and revised down to 11 after the first full run through. After the second run through, the categories were further adjusted to form six main categories and seven subcategories. These codes were then deductively applied to the L1 Arabic responses. As before, an additional round of coding took place in discussion with an L1 Arabic-speaking colleague to see whether 1) she agreed that the responses were appropriately grouped within the same category and 2) she would form a comparable label for that category. Through comparison and discussion, it was decided that a couple of the categories should be adjusted, but table 6.5 demonstrates the commonality between our category labels.

Table 6.5: Examples of category agreement between coders for Q13

My categories	L1 Arabic collaborator categories
Repeating	Repeating
…with focus on image	
…with focus on spelling	
Associating	Associating/connecting
…with letter shapes	…what they see and hear
…with familiar words/objects	…with real world/external knowledge
…with visual detail	
Remembering	Remembering

…specific sounds and letters	…first letter/sounds
…one aspect of the input	…sounds (perception/production)
	…images
Comparing/contrasting	Comparing/grouping
Evaluating learning	Evaluating
Directing learning	

Taking together the data from both language groups, six main strategy categories were formed, as well as seven subcategories, which are fully outlined in section 6.4.2. Coding guidelines were given to two additional coders for each language group, alongside a sample of 20% of the questionnaire responses[39]. The materials shared with coders can be found in appendix XVI and the R script used to calculate intercoder reliability with Krippendorf's alpha is available in appendix XVII.

6.4.2 QCA of language learning strategies

Every unit of analysis was assessed for its relevance to the category definition and coded with the level of abstraction in mind (see table 6.5). This meant that units of analysis must include reference to an action chosen and purposefully applied to the task of learning the (phonological) form of new words, with particular attention paid to the influence of modality of input. To focus on the active nature of strategy application, all categories were formed with the present participle of a verb (Oxford, 2016). As mentioned in section 6.4.1, the analysis resulted in the formation of six main categories of strategies, including seven further subcategories, which were considered in relation to broader cognitive and metacognitive strategies. A hierarchical diagram of the categories and subcategories is displayed in figure 6.1 and they are further defined in table 6.6. The definitions are given in English but contain both English and Arabic keywords and examples, to guide the coding of responses in both languages.

[39] The L1 Arabic coder who previously misunderstood the task was not invited back and a different L1 Arabic PhD researcher participated in the coding instead. The collaborative coding phase with the L1 Arabic data was completed in discussion with one of my L1 Arabic-speaking colleagues who contributed to the intercoder agreement coding in chapter 5. Meanwhile, the previous collaborative coding partner in chapter 5 swapped to code the 20% sample for intercoder agreement for the data in this chapter. This was a pragmatic decision made on the basis of the availability of my colleagues.

Figure 6.1: Hierarchy diagram showing the relationship between the general types of strategy, the main categories, and subcategories in response to Q13

Table 6.6: Coding guidelines for Q13 and LLS used during the task

Type	Category	Definition	Examples
Cognitive	Remembering	Broad focus on trying to remember the word, without specifying further. **English keywords** Remember, memorise, recall **Arabic keywords**: تذكر, حفظ, استرجاع [*remember, memorise, retrieval*]	“At this point my remembering” “بالاعتماد على الذاكرة فقط” *[relying only on memory]*
	Associating	Broad focus on associating different elements of the input, usually sound and image. It may include spelling and may be connected to memory. It does not specify further. **English keywords** Linking, connecting, relating, sounds like, looks like **Arabic keywords** إنشاء علاقة ,الربط [*create relationship, link*]	"link between what i hear and see” "اربط الصورة بالصوت و الكلمة" [*Associate the picture to the sound and the word*]
	With something familiar	Specific focus on associating new words with familiar words in a known language or an object. Can involve creating stories, contextual sentences or movement to facilitate the association.	“By sound associations with word I already know.”

	English keywords Looks like, sounds like, reminds me of, associate with, familiar **Arabic keywords** ربطها بكلمات اعرفها, تذكرني بشي اعرفه, مشابهة [*connect it to words I know, remember it with something I know or similar*]	"بربطها باشياء اخرى في الحياة اليومية مثلا كلمة fumel تبدا بحرف الفاء و تشبه الوردة " flower" " [*by linking it to other things in everyday life, for example, the word fumel begins with letter "fa" and looks like a flower "flower"*]
With visual detail	Specific focus on associating the word with detail from the image, e.g. colour, shape, texture etc. **English keywords** Colours, distinctive features, aspect of shape **Arabic keywords** بالألوان, خصائص مميزة, ناحية شكلية [*with colours, distinctive characteristics, form of the shape*]	"concentrating on the visual image and the colors in the image."
With letter shapes	Specific focus on associating the image with the shape of the first letter or shape of the written word. **English keywords** Looked like letter, make a mental shape **Arabic keywords** ربط الحروف بالشكل [*connect the letters with the shape*]	"Linking shapes for the letters with the object – eg Mackem had 3 legs, like an M" "الاعتماد على الذاكرة فقط وربط الحروف بالشكل" [*relying only on memory and connecting the letters to the shape*]

	Repeating	Broadly states the use of repetition out loud or in the mind/silently. It may reference memory but does not specify further. **English keywords** Repeat **Arabic keywords** أحاول نطق, التكرار [*attempt pronunciation, repeat*]	"I repeated them to myself out loud after hearing and or seeing them." "Repeated the word in my head" "أحاول نطق الكلمة" [*Trying to pronounce the words*] "بعض الكلمات ترسخت في ذاكرتي بسبب التكرار" [*some words stuck in my memory due to repetition*]
	Comparing/ contrasting	Looking for similarities and differences between sounds, letters, images and words. May include grouping or finding patterns. May refer to memory. **English keywords** Comparing, contrasting, differences **Arabic keywords** الكلمات التي تختلف, ملاحظة الإختلاف والتشابه والمقارنة [*words that differ, notice difference and similarities and compare*]	"I attempted to memorise the first letter which was often the difference between the two similar sounds eg the voiced vammell and voiceless fammell" "في مرحلة التعليم، كنت أحاول ربط الكلمات التي تختلف في صوت واحد فقط، مثلا كلمة "ناست" و"ماست" تختلفان في الصوت الأول فقط، وهذا سهل علي التعلم" *[In the learning phase, I was trying to associate words that differed in just one sound, for example the word "nasit" and "masit" only differed in the first sound, and this was easy to learn*]
	Visualising	Creating mental images/pictures based on the presented input. May be used to combine image,	"visualizing how the words sound incoproated with how they're spelt"

		spelling and/or audio in memory. **English keywords** Visualise, create/imagine picture/image, mind/mental. **Arabic keywords** التخيل /إنشاء صورة عقلية/ ذهنية [*imagine/ create a mental image*]	"Mental image"
Metacognitive	Focusing…		
	On sounds	Focus on remembering or paying attention to the sounds of a word, in terms of perception, recognition or production. **English keywords** remember/ memorise/ focus, sound/ voice/ pronunciation **Arabic keywords** تذكر/ حفظ/ التركيز على الصوت المسموع/الصوت/ النطق [*remember/ memorise/ focus on audible sounds/ voice/ pronunciation*]	"I tried to remember the voice saying it" "أحاول تذكر استماعي للكلمة" [*I tried to remember listening to the word*]
	On letters	Focus on remembering or paying attention to the letters or spelling of a word, usually the first letter. **English keywords** remember/ focus; first/last letter, spelling	"From remembering the first letter" "بالحرف الأول من كل كلمة" [*with the first letter of each word*]

	Arabic keywords تذكر/ حفظ/ التركيز على الحرف الأول/ الإملاء [*remember/ memorise/ focus on the first letter/ spelling*]	
On images	Focus on remembering or paying attention to the images, as a whole or a detail of the picture. **English keywords** concentrate/ focus/ study / remember, image/ picture/ object/ shape **Arabic keywords** تذكر/ حفظ/ التركيز الصورة /الشئ / الشكل [*remember/ memorise/ focus on the picture/ thing/ shape*]	"concentrating on the visual image and the colors in the image" "من خلال حفظ الصورة" [*by memorising the image*]
Evaluating learning	Evaluating success of learning, in general or specific strategies. May include switching strategies based on evaluation. **English keywords** Helped more, successful, better/worse **Arabic Keywords** ساعد أكثر, ناجحة, أفضل/أسوأ [*helps more, successful, better/worse*]	"I also sometimes used the shape of the item e.g. macum looked a bit like an "m" so I used that to remember. But definitely the repetition helped more than anything." "بمحاولة تذكر نطقها ولكن شعرت بتداخل بين كلمات علقت بذهني غير مرتبطة بالصور الصحيحه" [*by trying to remember the pronunciation, but I felt an overlap between words stuck in my mind that were not related to the correct picture*]

	Directing learning	Structuring the approach to the task by directing attention, going beyond the task, or choosing how to engage with input. Often mentioned in relation to cognitive strategies. **English keywords** Intentionally, rhythm, structure, game, focus **Arabic keywords** التركيز المتعمد, إيقاع, منظم, لعبة [*focus intentionally, rhythm, organise, game*]	"I said them out loud twice after each was shown, I then repeated the names once more before the testing phase. I almost had it in a rhythm that kept the learning time very fast-paced but consistent." "لم أركز كثيرا على الحروف المتشابهة كي لا أتشتت كنت أتعامل مع كل كلمة على حدى" [*I didn't focus much on the similar letters, so I wasn't distracted. I was dealing with each word individually*]

Intercoder agreement was acceptable across both language groups, with reliability measured as α = .71 between the L1 Arabic coders and α = .84 between the L1 English coders. This likely reflects the comparative quality of the data, as the L1 Arabic group mostly gave shorter and more ambiguous responses. Additionally, the areas of lower agreement often related to the subcategories within a common main category, where there was a higher level of agreement in interpreting the general meaning of the responses. The use of strategies in combination with each other also impeded clear coding of the data at times.

The full data set, including the final categories tags, can be found in appendix XVIII for both the L1 English and Arabic-speaker responses. Overall, 19 L1 Arabic responses were excluded, either because they did not answer the question or they did not mention a strategy that aligned with the category definition for the analysis. For example, some participants mentioned that they guessed[40] or did not really try to learn the words. This, therefore, did not meet the criteria of a strategy being purposely applied to the task with the goal of learning. A further four L1 English responses were excluded for this reason. Additionally, four L1 Arabic and one L1 English response indicated the main strategy used was making notes. These responses were excluded from further analysis on the basis that participants had been encouraged not to make notes. Additionally, there is agreement with Oxford (2016) that note-taking is complex and formed of multiple strategies, which are not detailed sufficiently in the responses to systematically examine. However, one example is described below, for the additional insight it offers to the overall discussion.

After participating, one L1 English-speaker contacted me to discuss the notes she made during the study and kindly allowed me to share them here. Figure 6.2 shows sketches and spellings for 10 out of 12 words from the learning phase. The notes demonstrate that this participant has been able to lexically encode the target contrasts for all the items, but chooses to denote /f/ with the digraph <ph> in two out of three instances, including the word where the English spelling was presented as <famis>. The use of a different vowel letter in the second syllable of /væməs/, compared to the presented spelling of <vamis> , also indicates that the participant drew on broader literacy experience and sound-spelling correspondences, rather than the written input provided. Scribbles show that a few of the

[40] Oxford (1996) mentions guessing in relation to compensation strategies. However, this involves drawing on additional context, which was not available in the present study.

words required multiple exposures to capture the final form, including evidence that /fæməs/ - /væməs/ were initially perceived as homophonous, even with the accompanying English spelling. The use of underlining also shows a focus on the second syllable of words. A final observation to consider is the fact that the images were novel items with low nameability. This is a common approach within psycholinguistic word learning studies (Carroll, 2013; Hayes-Harb & Barrios, 2021) and was successful in controlling participant exposure to items, as well as low confusion about the fact the target items were invented words. However, the ecological validity of this approach requires some attention. For example, when learning real words, it would be common for learners to write down associated L1 translations and use visual prompts, more so than the comparison of specific aspects of images, in the way that this experimental design potentially encourages.

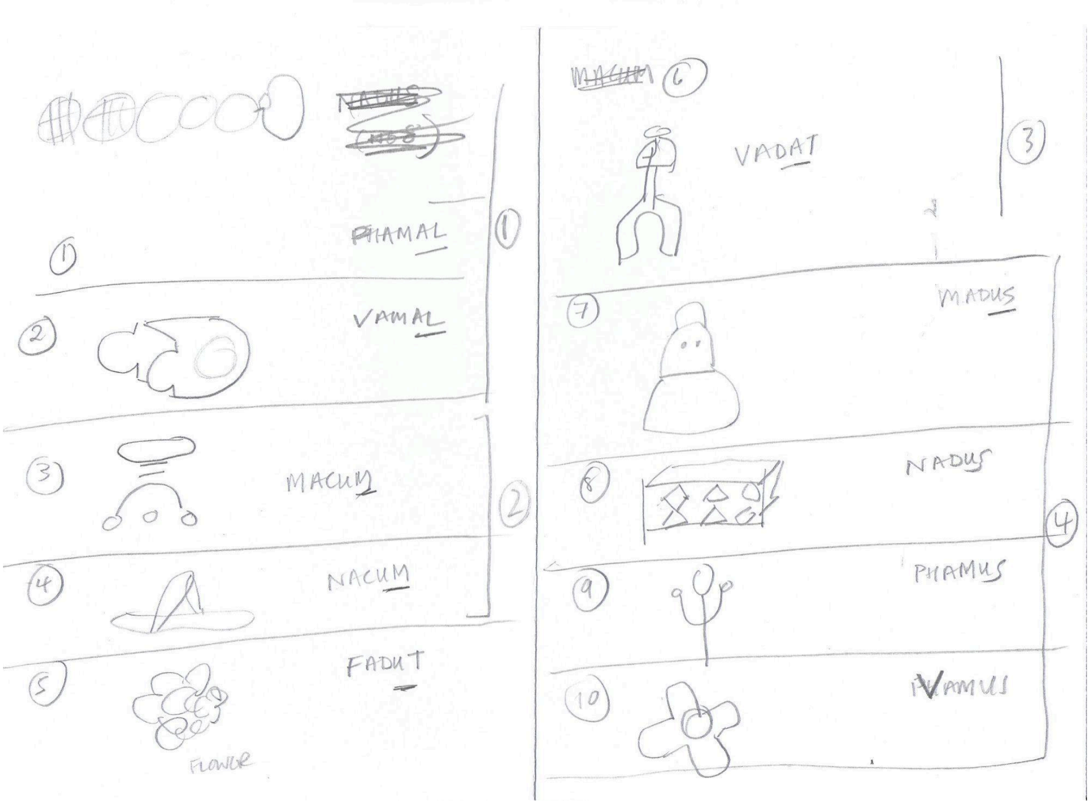


Figure 6.2: Notes made by one L1 English participant during the experiment

An alternative perspective is offered by Carroll (2013), who briefly discusses what counts as meaningful to learners in first exposure SLA studies, differentiating between reference and sense. Reference is understood to be the connection between language and non-linguistic objects, such as individuals, places, concepts, and properties, used in order to express both the experienced and imagined. In contrast, sense "is the semantic property that permits us to combine small meaningful units into larger meaningful units" (Carroll, 2013, p. 134). Thus, participants in the present study can be seen to be making sound-picture referential meanings when learning the words. The one-to-one mappings of sounds and pictures in these kinds of experiments do not reflect the correspondences between objects and words in natural languages, which often can be referred to by multiple words. However, the example is given of the way we readily learn proper nouns, such as people's names, demonstrating how people learn and retain referential mappings, even after minimal exposure. The way participants engaged in referential meaning-making is further exemplified and discussed in the next section.

As with the analysis outlined in chapter 5, the responses that were sufficient to be coded for specific strategy categories were analysed and presented in the results below. Each unit of analysis was a clear semantic unit, meaning that longer responses could be broken down into multiple units of analysis. The combination of fewer exclusions and longer responses meant there were more units of analysis for the L1 English group than the L1 Arabic group. The analysis below represents the data from 105 L1 Arabic and 164 L1 English coded units, with an overview provided in table 6.7.

Table 6.7: Overview of Q13 QCA results with number of analysis units for each category (% total units), by L1 group

Strategy category	Strategy subcategory	L1 Arabic	L1 English
Cognitive			
Associating		39 (37%)	73 (45%)
	broadly	*11 (10%)*	*20 (12%)*
	With something familiar	*21 (20%)*	*32 (20%)*
	With letter shapes	*3 (3%)*	*12 (7%)*
	With visual detail	*4 (4%)*	*9 (5%)*
Repeating		5 (5%)	28 (17%)
Remembering		8 (8%)	3 (2%)
Visualising		3 (3%)	5 (3%)
Comparing/contrasting		1 (1%)	6 (4%)
Metacognitive			
Focusing		44 (42%)	42 (26%)
	On sounds	*32 (30%)*	*18 (11%)*

	On letters	*6 (6%)*	*10 (6%)*
	On images	*6 (6%)*	*14 (9%)*
Evaluating learning		3 (3%)	4 (2%)
Directing learning		2 (2%)	3 (2%)

6.4.2.1 Cognitive strategies

Cognitive strategies are understood to relate to mental processes and memory, whereby learners consciously think and behave in a way to facilitate comprehension, learning and retention of information. Participants were not directly asked whether they perceived the reported strategies to be reflective of cognitive processes, as opposed to other domains. Rather, the grouping of these strategy categories reflects the interpretation of responses and understanding of this construct within the relevant literature (see chapter 2.1.3.3 and section 6.2). Within this analysis *associating*, *repeating*, *remembering*, *visualising*, and *comparing/ contrasting* strategies are interpreted to reflect cognitive strategies. The majority of responses, across both language groups, make reference to one or more of these cognitive strategies, specifically 56 (53%) L1 Arabic and 115 (70%) L1 English response units.

Associating

When asked what strategies were used to learn the words during the experimental task, 39 (37%) L1 Arabic and 73 (45%) L1 English units indicated that participants' often attempted to learn the words by making associations. Within this category, 11 L1 Arabic and 20 L1 English units do not specify beyond broadly associating aspects of the input. For example, L1 Arabic responses make general references to "connecting" (ربط), "linking" or "relating" words with the pictures. Similar wording was found in the L1 English group, as well as terms like, "applying", "associating", "matching", and "grafting". Quotes included here shed further light on the influence of orthography when associating. For example, EN5649032 draws on existing literacy experience to filter sounds through possible letter representations. This comment implies a conscious process of assimilation to an L1 category, as represented by L1 grapheme-phoneme correspondences.

By making connections between the sound of the word (more than its spelling - but thinking about what letters I thought I was hearing in it) and the image of the object the word described. (EN5649032)

Responses often highlight the creativity and idiosyncrasy of associations, including the use of stories, songs, and actions. In the subcategory of *associating with something familiar*, which includes 21 L1 Arabic and 32 L1 English units, participants draw on the languages they are familiar with to make associations with similar sounding words, which usually resemble the object in some way. For context, a reminder of the stimuli is presented in figure 6.3. Responses often include example words, demonstrating L1 transfer and the drive to connect the word forms to semantic content. The examples below focus on words differing by /f-v/, due to the interest in how L1 Arabic participants navigate the difficulty of this contrast. Interestingly, none of the L1 Arabic responses include target items beginning with /v/, except from AR5336722. Even this participant associates both "famo" and "vamo" with "فم", meaning "mouth" in Arabic. This example offers evidence of category assimilation through L1 transfer, despite perceiving the /f-v/ distinction.

Connecting the shape to something in real world. e.g. those that are called 'famo' or 'vamo', I try to connect the word heard which is a bit similar to "فم" in Arabic saying the shape of the those pictures can be entered inside the 'mouth=فم'. Another example, the shape that's called 'famis' ...I say it looks like a unique flower..so I remind myself that it's famis "like or driven from the English word famous". So in short, try to connect the new words to known words in either language Arabic or English. (AR5336722)

The connection between the invented word /faməs/ and the English word "famous" is noted in both L1 Arabic and English responses. While not explicitly mentioned, it is of interest that /faməs/ was presented with English spelling, perhaps enhancing the association to an existing English word starting with the same letter. The focus on the first letter only and the association with the picture by AR5003023, shows that this strategy may have led to confusion between /faməs/ and /faməl/. Additionally, the script switching in these examples indicates that, even without exposure to OI, words were perceived to be English and lexically encoded with English orthographic information.

ربط شكل اللعبة مع الكلمة. مثلا كلمة فامس شكل اللعبة شخص رافع يديه للأعلى. فامس من فيموس تعني مشهور و هذا المشهور سعيد و رافع يديه (AR4986647)

[*Connect the shape of the toy with the word. For example the word "famis" toy shape is a person who holds his hands up. "Famis" from "faymus" means famous and this famous person is happy and holds up his hands*]

بربطها باشياء اخرى في الحياة اليومية مثلا كلمة fumel تبدا بحرف الفاء و تشبه الوردة "flower" (AR5003023)

[*by connecting it with something else in daily life like the word "fumel" begins with the letter "fa" and looks like a flower*]

ربط الكلمة بالشكل حتى لو بشكل بعيد جدا. مثلا famis شكلها يشبه الجهاز التناسلي للأنثى —>female famis. و nackle تشبه مخدة الرقبة neck. (AR5442543)

[*connecting the word to the shape even if it is very far from the word. For example famis's shape resembled the female reproductive system -> female famist. And nackle resembles a neck pillow - neck*]

Figure 6.3: Overview of target words, with accompanying images and written form

The L1 English responses demonstrate similar connections are made between items and words in English e.g., famis -> famous person holding up their hands (EN5312548). Additionally, they draw on other languages they know to make connections, such as French and German. In these responses there is some evidence that, in contrast to the L1 Arabic group, the L1 English-speakers do not assimilate /f-v/ contrast items to a single category. However, there are still more examples of /f/ initial words in general. Across both groups, participants search for familiar connections between the target items, which can take the form of similar objects, starting sounds, and general overlap with a known word.

> Usually with some kind of mental association (e.g. 'famis' sounds like 'famous' and the picture looked like a person raising their arms in triumph) and repetition out loud. (EN5312548)
>
> I tried to associate the picture with an english word that reminded me of that sound e.g. the black object looked mechanical so i thought of 'v' for vaccum, the flower object started with an 'f', the massid one looked like a mess! (EN5366758)
>
> I thought about what the pictures looked like and tried to create associations even if they were weird. eg I think it was 'famil' that I remembered as lots of little loops making a family. Or for 'nackam' I thought about a back being 'knackered' and that it looked a bit like someone lying on their back. (EN5758767)

Another subcategory was formed to reflect the strategy of *associating with letter shapes*. There are three L1 Arabic and 12 L1 English units in this subcategory, which mostly include examples of visualising the initial English grapheme within the shape of the object. As the examples mention items presented with audio only or with Arabic spelling, the connection to English letters is not dependent on English spelling exposure. Therefore, this exemplifies another way in which participants draw on their existing literacy experience, across languages and scripts, to remember the words. It is also notable that, again, L1 English responses are the only ones that reference /v/ initial words. Across both language groups, participants use combinations of *associating* strategies alongside other strategies, such as *repeating* and *comparing/contrasting*. For example, in the quotes below, the units relevant to the subcategory *associating with letter shapes* have been highlighted in bold but kept in the context of the full response to show strategy combinations.

> I linkd the shape of the object with a word or letter. **for example the last shape in the previous question looks like M**. The Madas shape remined me with "madas" which means shoes in Makkah & Jeddah people slang. there alos an object looks like fallopian tube. (AR5341799)
>
> i paid particular attention to the beginning letter as this is what was varying between the new words. **I also sometimes used the shape of the item eg macum looked a bit like an "m" so I used that to remember.** But definitely the repetition helped more than anything. (EN5336201)
>
> There are some pictures that I tried to match with something like **there was a vedet? which was a picture of something that looked like an upside down V to me?** And then the picture that looked like a sandal was near the beginning and I thought it sounded like nechem...that one seemed to stick in my brain easier, I don't know why? (EN5595978)

The final subcategory is *associating with visual detail*, which includes four L1 Arabic and nine L1 English units. A few examples are provided below; however, discussion is limited as these strategies mostly demonstrate memorisation processes related to the artificiality of the task. The fact that one participant identifies this strategy as a “shortcut” further indicates that this is not applied with the intention of accurately learning the word form. Mostly, participants in both groups focused on distinctive colours and shapes within the object, often in combination with repeating.

> sometimes repeating the word but not always(this failed me) + creating an image for the pic+word... ex the word famis for me seemed like a famous person who held his\her hand high.. trying possibly to wave to the fans? other than that, **some images were easier to remember because of the color(i.e., the orange thingy)**. (AR5336140)

I mouthed or said them aloud as I went through. I also sometimes made a mental note of a **distinctive feature of the picture, like 'blue blob' or 'upside-down yellow Y'** (EN5503474)

I found 'Mackem' the easiest to remember, as this is a real English word for someone from (or the dialect of) Sunderland / Wearside. For everything else, **I tried to associate the word with the colour of the object rather than the object itself as a shortcut**, **but this got hard when there were three orange objects.** (EN5998960)

Repeating

The next category of *repeating* includes five (5%) L1 Arabic and 28 (17%) L1 English units. This strategy is mostly applied alongside other strategies, such as *focusing* and *associating*. Some participants specify that they repeat the words aloud, whilst others repeat silently/in their heads. Additionally, some mention mouthing the words, repeating them multiple times, or timing their repetitions at specific moments of the task. As noted by EN5336201, some responses evaluate this as a successful strategy, while others, like AR5336140, find it to be unreliable, both quoted above in relation to *associating*. The examples below show the relationship between repetition and memory, as well as combinations with other strategies.

بعض الكلمات ترسخت في ذاكرتي بسبب التكرار (AR5341797)

[some words firmly set in my memory due to repetition]

studying the picture whilst **repeating the word over and over** (EN5587514)

I repeated them in my head when the picture was shown. I also tried to link the words to the images in my head. (EN5690497)

Remembering, visualising and comparing

Other cognitive strategies include *remembering*, *visualising,* and *comparing/contrasting*. There are fewer instances of these categories and references tend to be quite vague. Regarding *remembering*, eight L1 Arabic and three L1 English units do not specify beyond ambiguously trying to remember the words or pictures. Similarly, only three L1 Arabic and five L1 English units mention visualising, picturing, or creating a mental image of the word, usually combined with the object and spelling. Some responses imply creating imagined connections or using this visualisation to look at patterns between the words, tying in with looking for similarities across the words. Only one L1 Arabic response makes reference to *comparing/ contrasting* elements in the input, whereas six L1 English units are included in this category. Some responses indicate that participants figured out the experimental aim and thus apply metacognitive strategies to direct their focus to learning the minimal pairs,

> في مرحلة التعليم، كنت أحاول ربط الكلمات التي تختلف في صوت واحد فقط، مثلا كلمة "ناست" و"ماست" تختلفان في الصوت الأول فقط، وهذا سهل علي التعلم (AR5370516)
>
> *[In the learning stage, I was trying to associate words that differed in just one sound, for example the word "nasit" and "masit" only differed in first sound and that's easy to learn]*
>
> focusing on the consonants, especially those at the beginning of each syllable. **I also realized you were mostly using words with minimal pairs (f/v, m/n) so focused on those especially**. (EN5751031)

6.4.2.2 Metacognitive strategies

Metacognitive strategies refer to processes which manage, evaluate and organise learning, examples of which are given in chapter 2.1.3.3 and section 6.2. These strategies are typically found in combination with cognitive strategies, which they then regulate and monitor. As in the previous section 6.4.2.1, this grouping of categories reflects theoretical understandings of the construct in relation to participant responses. Within this analysis *focusing*, *evaluating learning,* and *directing learning* are interpreted to reflect metacognitive strategies. Both language groups make reference to these metacognitive strategies, specifically 49 (47%) L1 Arabic and 49 (30%) L1 English responses.

Focusing

The largest category within metacognitive strategies was *focusing,* including 44 (42%) L1 Arabic and 42 (26%) L1 English units. There may appear to be overlap with the category of *associating*. However, the separate categories reflect the intention to stay close to the wording of participants and the apparent distinction between focusing on one, or multiple, aspects of the input, compared to actively creating associations. Also, in contrast to *associating*, participants tended to provide shorter and more ambiguous responses in this category. As such, this category is interpreted to display the tendency of participants to focus/concentrate on or pay attention to specific elements of the input, individually or in combination, to remember the target items. Furthermore, these responses imply a management of the input and other strategies to learn the new words.

For most L1 Arabic responses (32 units), this involved *focusing on sounds*. Most participants mentioned recalling what they heard, frequently mentioning the sound (صوت) and pronunciation (نطق). A couple of responses added details, such as breaking the word down into syllables, focusing on initial sounds, and deliberately covering written forms. These responses demonstrate that at least 30% of L1 Arabic-speakers were intentionally learning the phonological form of target items and were consciously attending to pronunciation. There is also evidence of written forms distracting from the auditory input and difficulty suppressing attention. Hence, the need to employ a further metacognitive strategy of *directing learning* and physically covering the spelling (AR5730442). Another point to note, as mentioned in chapter 5, is that participants struggle to delineate spelling and sound (AR5037277). The 18 L1 English responses in this subcategory also make frequent mention of remembering particular sounds, sublexical analysis by focusing on syllables and beginning consonants (EN5751031), and the conflation of sound and spelling (EN5359677).

> بمحاولة تذكر نطقها (AR5010884)
>
> [*I tried to remember her pronunciation*]
>
> ركزت على الصوت، في بعض الأحيان كنت أتعمد تغطية الكتابة بأصابعي (AR5730442)
>
> [*I focused on the sound, sometimes I deliberately covered the writing with my fingers*]
>
> I was trying to remember the spelling.. I mean the sound 🔉 (AR5037277)
>
> Focusing on the consonants, especially those at the beginning of each syllable. (EN5751031)
>
> Distinctive sounds, especially first letter (EN5359677)

The subcategory of *focusing on letters* contains six L1 Arabic and 10 L1 English units. Participants in both language groups focus on the first letter in order to remember the words, disproving a hypothesis presented in chapter 5 that L1 English and L1 Arabic participants may differ in this regard. It is again worth mentioning that, due to the common conflation of sounds and letters, it remains unclear whether these responses refer to written input, existing literacy experience, or refer to sounds rather than letters. However, AR5027570 makes it clear that they focus on pronunciation and then speculate the spelling based on their language and literacy experience.

بالحرف الأول من كل كلمة (AR5341802)

[*with the first letter from each word*]

i tried to think about how that person say it and guessed the spelling upon my experience (AR5027570)

i focused on the first letter (EN5659195)

Rembering the first letter of the word and placing it with the colour of the object in my memory (EN5642980)

The example above from EN5642980 demonstrates how *focusing* strategies can be combined with *associating with visual detail*. The units in the subcategory of *focusing on images* then reveal that some only mention focusing on the shapes and colours of the object, which does not imply intentional learning of the phonological forms of the target items. However, of the six L1 Arabic and 14 L1 English units in this subcategory, *focusing on images* is often mentioned in combination with *repeating* or *focusing on sounds*, reflecting referential pairings. The example from EN5719579 also indicates that strategies can vary depending on the input provided. Thus, in contrast to AR5027570, visualisation of the spelling with the image was not speculatively imagined based on existing literacy experience, but rather depended on what was available in the input.

من خلال حفظ الصورة (AR5622524)

[*Through memorising the picture*]

عن طريق التركيز فى الصوت والصورة (AR5553283)

[*By **focusing on** the sound and **image***]

Look at the shapes/colors of the objects and say the words to myself. If spelling included, in English, I included that with my memory of the object. (EN5719579)

Evaluating and directing learning

Two other metacognitive strategies are referenced in a few participant responses, namely *evaluating learning* and *directing learning*. Some examples have already been given of these, as they are always mentioned alongside other strategies. Three L1 Arabic units and four L1 English units show that participants evaluate how successful their strategy usage is or how well they have learned the words. Meanwhile, two L1 Arabic and three L1 English units demonstrate participants direct their learning, by creating additional ways to approach the task, testing their memory, or managing their pace of learning. For example, some mention closing their eyes or using their hand to hide parts of the presentation, as well as trying to name all four objects when they appeared in the learning trials and not just select the one matching the audio. The fact that some participants feel it is beneficial to limit their exposure to the visual input, particularly by covering the spelling, demonstrates how an awareness of orthographic influence on learning can be reflected in strategy use and management.

6.4.3 QCA strategy summary

Responses to Q13, asking participants how they tried to learn the words during the study, reveal that a range of different strategies are applied to the task. Figures 6.4 and 6.5 display the proportion of response units that are included in each identified strategy category and subcategory, by language background. These treemaps indicate that both groups predominantly reported focusing on different aspects of the input or making associations to better remember the new words. The L1 Arabic participants frequently state that their primary strategy was *focusing on sounds*, where some break words down into smaller units to attend to individual sounds and syllables, while others concentrate on the sound of the whole word. Due to the short responses, it is difficult to speculate about phonological strategies beyond this, but it is interesting to note that this strategy is often combined with both *focusing on the images* and *letters*. In these cases, participants are conscious of different weight being given to cues in the input, with frequent mention of salient features, such as distinctive sounds, first letters, and particular colours. Responses which mention *focusing* strategies also conflate sounds and spelling, draw on literacy skills to guess the spelling of words, and vary in either attending to or suppressing distraction from written input, sometimes in combination with other metacognitive strategies, such as *directing learning*.

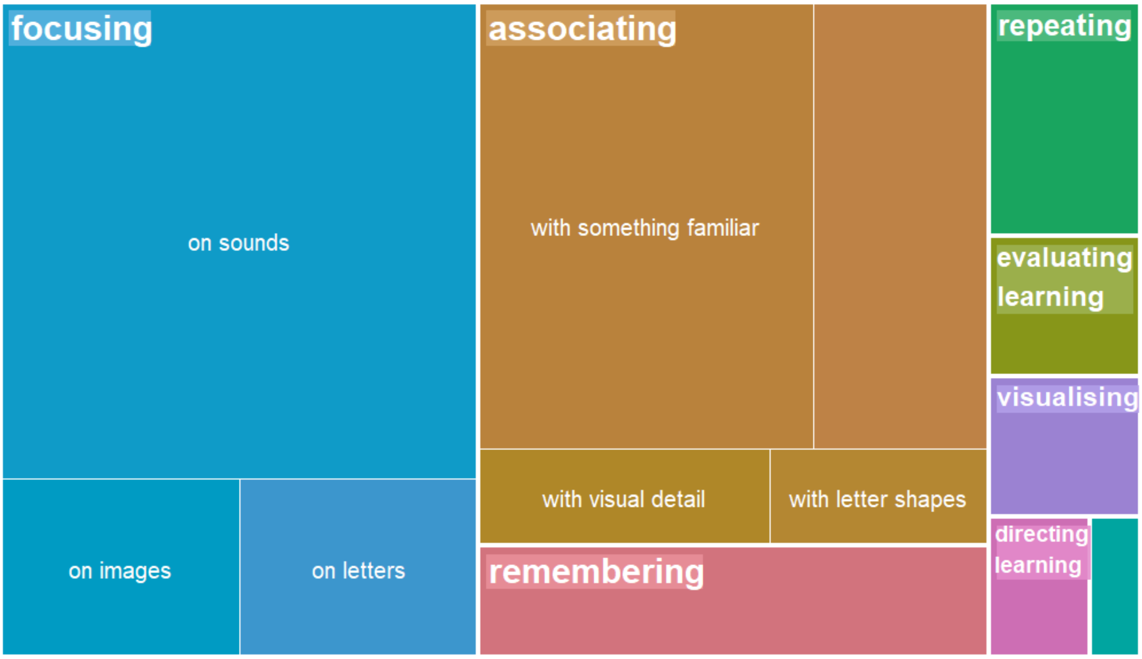


Figure 6.4: Treemap of L1 Arabic group's self-reported strategies during the study

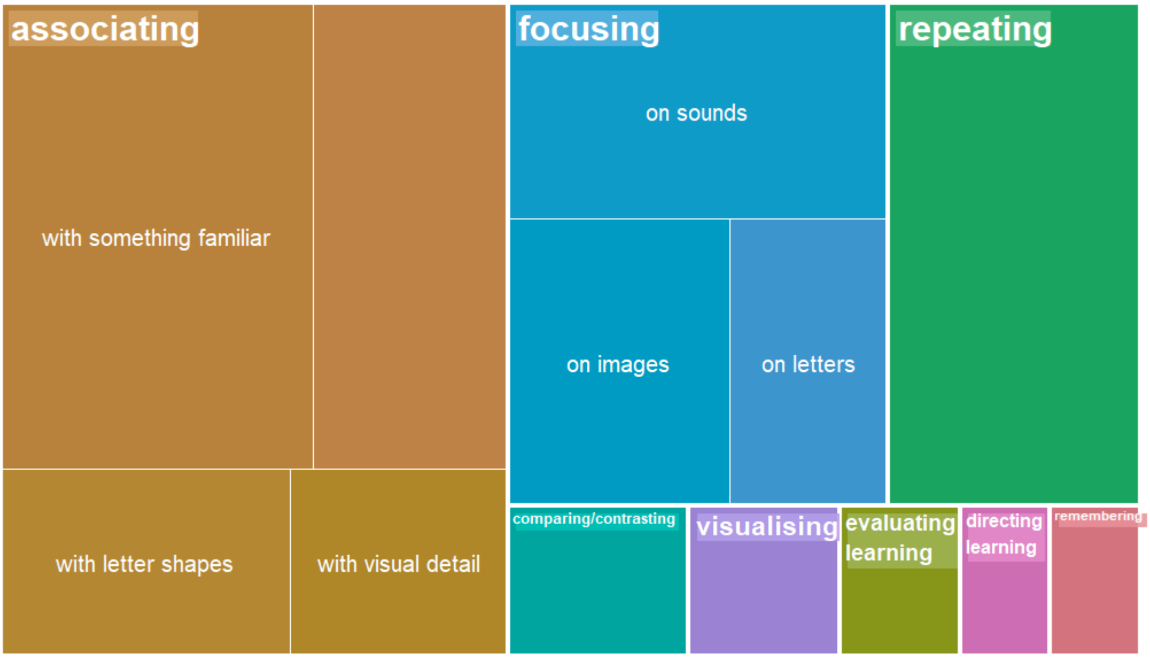


Figure 6.5: Treemap of L1 English group's self-reported strategies during the study

The emphasis that participants place on trying to remember aspects of the images, mentioned across *focusing* and *associating* strategies, highlights methodological issues around the use of unusual objects with low nameability. It is evident from several responses that participants are reporting strategies for quick memorisation and not reflective of a goal to learn the word form. This would benefit from development if a similar procedure was applied in future research. Perhaps, the original multi-session design (see chapter 3.2.1) could offer some possibilities to extend the investigation further into the territory of learning, or a more participative design involving interaction with similar but tangible objects. That being said, responses mentioned in this and the previous chapter reflect processes and goals that align with language learning and referential meaning-making. Therefore, these findings work to shed light on the stated research questions but also reveal the need for, and direction of, further enquiry.

The detailed examples that participants provide in relation to creating associations are of particular interest in relation to cross-linguistic transfer and the use of orthographic knowledge to connect to the images in an unexpected way. Both groups predominantly mention *associating* in relation to familiar words, objects, and languages. Focusing on /f-v/ contrast items, there is only one instance of a /v/ initial word being mentioned in L1 Arabic responses, which is assimilated to the same L1 word as its /f/ initial minimal pair. The absence of /v/ initial items, and the explicit example of assimilation, suggest that this phonological contrast is not well-established in this group, compared to /m-n/ words or the L1 English group. It appears that strategies, such as associating, do not facilitate the lexical encoding of this confusable contrast, which is unsurprising due to clear cross-linguistic transfer. It is also noted that *associating with letter shapes* is another example of drawing on existing literacy experience to remember the words, particularly in English.

As documented by Oxford (2016), strategies are often used in sequences, chains, and clusters. This was the case within the present data set, where participants combine different approaches to *focusing* and *associating*, as well as adding strategies of *repeating*, *visualising*, and *comparing*. While the metacognitive strategy of focusing was mentioned most often, a wider range of cognitive strategies were reported in the data. Additionally, as would be expected, most metacognitive strategies were mentioned alongside cognitive strategies, demonstrating the conscious direction of attention and reflection to manage learning. Overall, the focus on cognitive and metacognitive strategies, rather than affective, social, or motivational strategies, aligns with previous studies exploring pronunciation strategies (Całka, 2011; Rokoszewska, 2012).

Returning to the research question at the start of this chapter, participants report a range of LLS in relation to the task. These strategies often make use of any existing phonological and orthographic knowledge to learn the new words, as well as targeting phonological and orthographic content in the input. It may seem obvious, but this is important to highlight, as studies exploring phonology and pronunciation strategies rarely refer to orthographic influence, even though letter knowledge appears to be a key point of connection between the phonological form and associated meanings for many in this study. Indeed, participants often struggle to separate sounds from letters in their reflections. The tight connection between phonology, orthography and semantic content in responses also echoes theoretical and empirical research into developing L2 lexical representations, as well as written and spoken language processing (see chapter 2.2). While the results presented here give an indication of the strategies participants remember applying to the task, it cannot speak to frequency of usage, the relationship between strategy use and performance, or how reported strategies relate to different learner characteristics. In order to triangulate results and address this aspect of the research question, complementary quantitative analysis is presented in the next section.

6.5 Quantitative Analysis

6.5.1 Descriptive overview of strategy frequencies

To further understand the role of LLS in the present study, Likert-scale responses to the POLLS inventory are explored to establish what strategies participants use and how frequently. An overview of mean frequencies for each item is found in table 6.8, and visualised in figures 6.6 to 6.8, where responses have been recoded in numeric form, i.e. “Never” = 1, “almost never” = 2, “sometimes” = 3, “almost always” = 4, “always” = 5 (Rokoszewska, 2012). The table is divided into cognitive, metacognitive and affective strategies, based on the strategy inventories which inspired the present study (Berkil, 2008; Oxford, 1990; Peterson, 1997; Rokoszewska, 2012). Frequent strategies, where the mean exceeds 3 (i.e., more than sometimes), are then highlighted in bold. This reveals that both the L1 Arabic and L1 English groups report frequent usage of 11 strategies, although the particular strategies and frequencies do somewhat vary between the two groups.

Table 6.8: Overview of mean strategy use frequency during the study (*SDs*)

Strategy	Questionnaire item	L1 Arabic (n=109)[1]	L1 English (n=117)
Cognitive	14) I created association with words or things I already know.	**3.5[2] (1.2)[3]**	**3.1 (1.3)**
	15) I put the word in a context to remember it (e.g. a sentence, a story, a rhyme).	2.4 (1.3)	1.8 (1.1)
	16) I made a mental image or imagined additional connections to help me remember.	**3.6 (1.2)**	**3.6 (1.1)**
	17) I visualised the spelling of the words in Arabic in my mind.	2.3 (1.3)	1.2 (0.5)
	18) I visualised the spelling of the word in English in my mind.	**3.7 (1.2)**	2.8 (1.2)
	19) I repeated the words out loud or in my head.	**3.4 (1.4)**	**3.8 (1.3)**
	20) I thought about or practised mouth positions to pronounce the words.	**3.3 (1.3)**	2.0 (1.2)
	21) I thought about an action or movement to help remember the words.	2.7 (1.3)	1.5 (0.9)
	21) I broke words down into syllables or sounds.	2.2 (1.2)	2.4 (1.3)
	22) I tried to find patterns in the new words or sounds.	2.8 (1.3)	**3.1 (1.3)**
	23) I tried to connect sounds and letters.	**3.4 (1.3)**	**3.1 (1.3)**
	24) I tried to remember the first or last sounds.	**3.5 (1.2)**	**3.7 (1.3)**
	25) I looked for similarities and contrasts between the pronunciation of the words and the languages I know.	2.7 (1.4)	2.8 (1.5)
	26) I grouped similar sounding words.	2.9 (1.2)	2.6 (1.3)
	27) I used the English spelling to distinguish between similar sounds.	2.8 (1.3)	**3.2 (1.4)**
	28) I used the Arabic spelling to distinguish between similar sounds.	2.1 (1.2)	1.1 (0.3)
Metacognitive	29) I purposefully ignored the English spelling.	2.0 (1.1)	1.7 (1.0)
	30) I purposefully ignored the Arabic spelling.	2.6 (1.4)	**3.5 (1.7)**
	31) I found ways to test my memory and recall the new words.	**3.3 (1.1)**	2.8 (1.1)
	32) I thought about my progress in learning the new words.	**3.3 (1.1)**	**3.1 (1.2)**

	33) I noticed my mistakes and used that information to help me do better.	**3.6 (1.0)**	**3.6 (1.0)**
Affective	34) I noticed I felt more relaxed or confident when I saw the Arabic spelling.	2.7 (1.3)	1.2 (0.5)
	35) I noticed I felt more relaxed or confident when I saw the English spelling.	**3.8 (1.1)**	**3.3 (1.4)**

[1] The number of participants is lower for the L1 Arabic group, as five participants did not respond to the inventory in the debrief questionnaire.
[2] "Never" = 1, "almost never" = 2, "sometimes" = 3, "almost always" = 4, "always" = 5
[3] Mean frequencies > 3 are in bold to show higher frequency items.

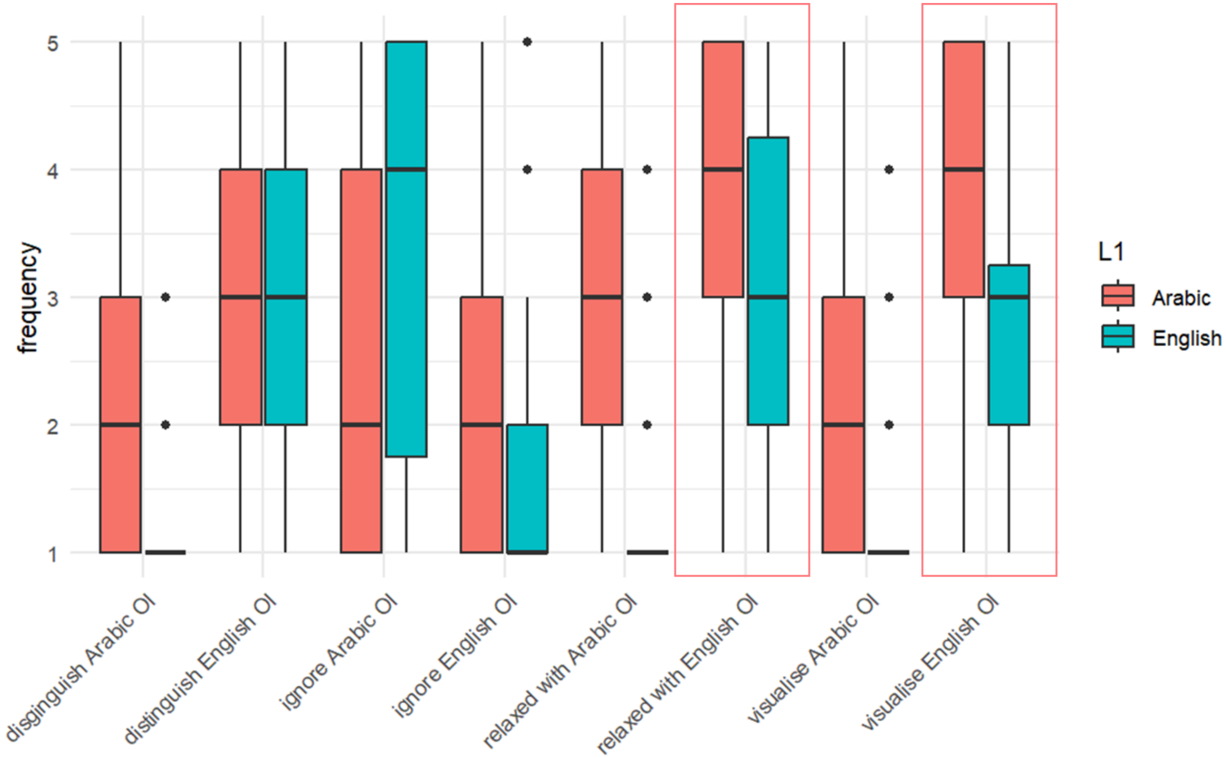


Figure 6.6: Boxplots for frequency of strategy use (scale 1-5) related to Arabic and English spelling, by L1 group. Frames highlight where the L1 Arabic group report high frequency strategy use.

The descriptive frequencies in table 6.8 demonstrate differing responses across the two language groups, which is particularly notable in relation to types of written input (figure 6.6). For example, the L1 English group never uses Arabic OI to *distinguish between similar sounds, feel more relaxed/confident* or *visualise the spelling* in the task, and almost always *ignore the Arabic spelling*. Meanwhile, the L1 Arabic group reported high frequency usage of English OI to *visualise the spelling, feel more relaxed/confident*, and almost never *ignore the English* or *Arabic spelling*. As highlighted in figure 6.6, the frequency of using English OI to *visualise the spelling, feel more relaxed/confident* is higher in the L1 Arabic group than the L1 English group, revealing a heavier reliance on English script input. Thus, strategy usage appears to differ in relation to experience with distinct script input and by language background. When entirely unfamiliar, distinct written forms are mostly ignored by the L1 English group. In contrast, familiar OI in a distinct script is often used as a tool for learning by the L1 Arabic group. It is also notable that *using Arabic* and *English written input to distinguish similar sounds* is comparable across both language groups and is not the most frequent orthographic strategy.

The boxplots in figure 6.7 illustrate that responses are broadly comparable across both language groups in relation to cognitive strategies, such as frequent usage of *remembering the first sound*, *grouping, creating mental images, seeking patterns, finding similarities with the L1*, and *connecting sounds and letters*. In contrast, both groups report infrequent use of *creating context* and *breaking words down into sublexical units*. One of the most frequent strategies for both groups is *creating associations*, echoing findings in section 6.4. Creating semantic connections through association, actions and context is more frequent in the L1 Arabic group than the L1 English group. Additionally, while both groups report high frequency use of *repeating* strategies, the L1 Arabic group reports *mouthing* the words more frequently than the L1 English group. Finally, both groups report comparable usage of metacognitive strategies, where *noticing mistakes* to improve performance is the most frequently used of the three strategies by both groups (figure 6.8). Overall, L1 Arabic participants report higher frequency usage of both cognitive and metacognitive strategies than the L1 English group.

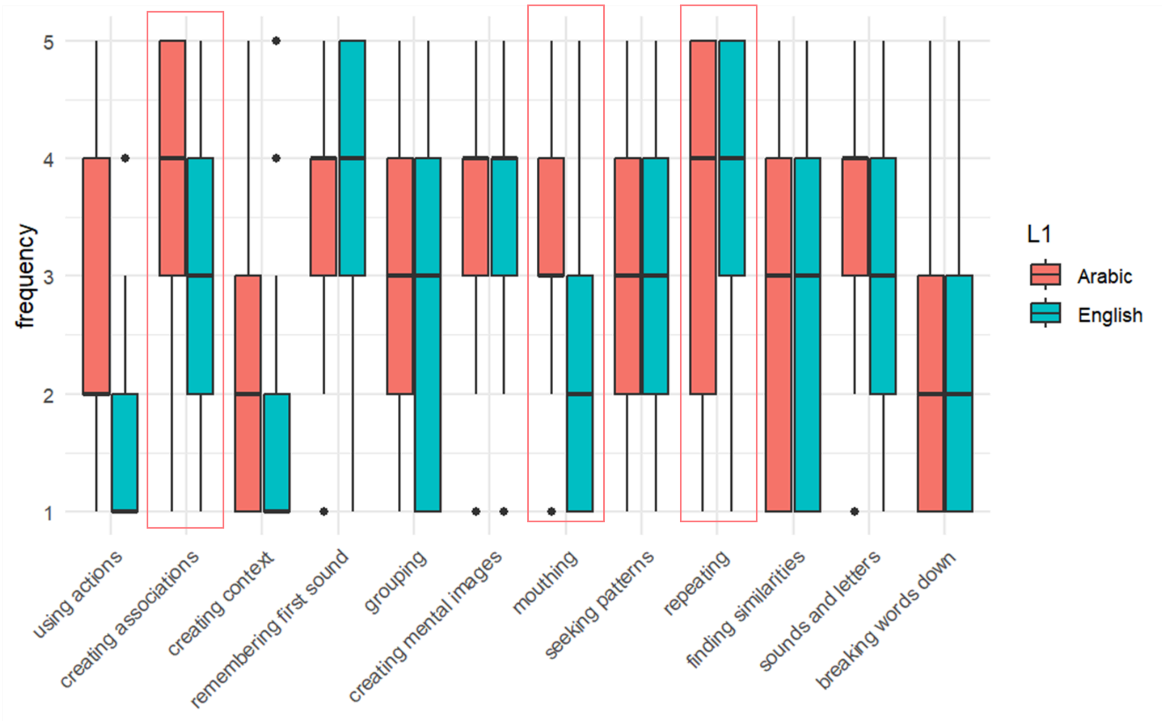


Figure 6.7: Boxplots for frequency of cognitive strategy use (scale 1-5), by L1 group. Frames highlight where the L1 Arabic group report high frequency strategy use.

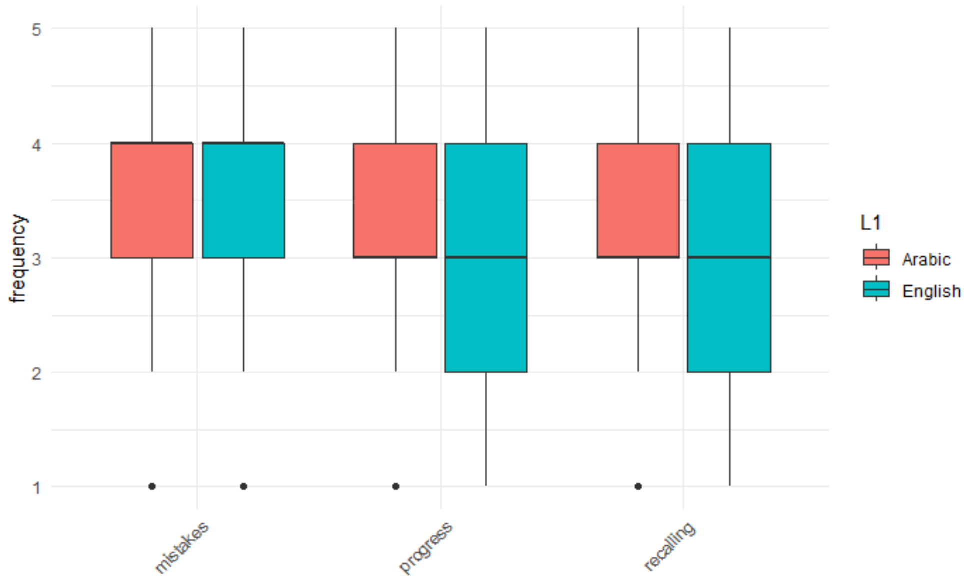


Figure 6.8: Boxplots for frequency of metacognitive strategy use (scale 1-5), by L1 group

It is well documented that proficiency is a relevant consideration both in terms of amount and type of strategies used by learners, where higher level learners tend to report using more strategies, as well as different types of strategies compared to lower level students (Griffiths, 2018). Table 6.9 provides an overview of the mean frequencies of strategy usage by self-reported L2 English level[41] of the L1 Arabic group. A closer look at these results by English proficiency sheds light on L1 Arabic strategy usage, again with particular focus on the role of orthographic input. For example, reliance on L1 Arabic spelling to *distinguish sounds* and *visualise words* declines with proficiency, while using L2 English spelling to *feel more relaxed/confident* increases with proficiency (figure 6.9). This could indicate a difference between *base* and *plus* strategies, favoured by beginners and advanced learners respectively (Griffiths, 2018). At each proficiency level, learners report frequently *visualising words in English* and they rarely *ignore English spelling*. The consistent use of orthographic strategies related to L2 English spelling suggests that use of L2 written input is a *core* strategy (Griffiths, 2018). Returning to the affective influence of orthographic input, even though there is an increase with proficiency, it is noteworthy that participants at all proficiency levels frequently draw on English spelling to feel more confident or relaxed about learning the new words.

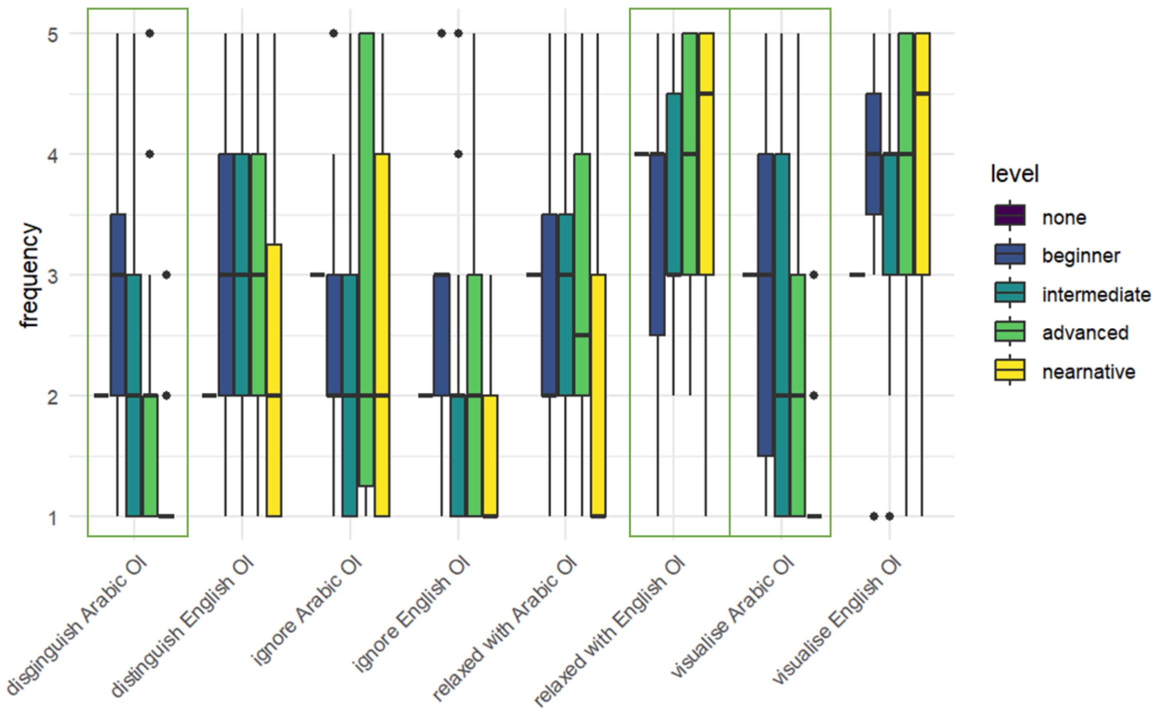


Figure 6.9: Boxplots for frequency of strategy use (scale 1-5) related to Arabic and English spelling, by proficiency. Frames highlight strategy usage that shifts as proficiency increases.

[41] The self-reported levels are used here for ease of data grouping. Subsequent analyses use the test scores from the short English test, as a more objective, continuous measure.

Table 6.9: Mean frequencies of strategy use by L2 English level of L1 Arabic group (*SDs*)

Strategy	Questionnaire item	Begin (n=11)	Inter (n=35)	Adv (n=50)	Near-nat (n=12)
Cognitive	14) …associations...	2.8 (1.2)	**3.2 (1.2)**	**4.0 (1.1)**	**3.5 (1.4)**
	15) …context...	2.0 (0.8)	2.2 (1.0)	2.6 (1.4)	2.4 (1.4)
	16) …mental image....	**3.2 (1.5)**	**3.4 (1.1)**	**3.7 (1.1)**	**3.5 (1.2)**
	17) …visualised the spelling in Arabic...	2.8 (1.5)	2.6 (1.3)	2.3 (1.3)	1.3 (0.6)
	18) …visualised the spelling in English...	**3.6 (1.4)**	**3.5 (0.9)**	**3.7 (1.2)**	**3.8 (1.4)**
	19) …repeated…	**3.6 (1.2)**	**3.7 (1.2)**	**3.3 (1.4)**	2.5 (1.6)
	20) …mouth positions...	**3.7 (1.3)**	**3.5 (1.3)**	**3.3 (1.4)**	2.6 (1.4)
	21) …action...	2.8 (1.6)	2.7 (1.2)	2.6 (1.2)	2.5 (1.5)
	21) …broke words down...	2.7 (1.3)	2.2 (1.1)	2.1 (1.2)	2.3 (1.5)
	22) …patterns...	2.5 (1.2)	2.5 (1.1)	3.0 (1.4)	**3.2 (1.5)**
	23) …connect sounds and letters...	**3.5 (0.9)**	**3.5 (0.9)**	**3.3 (1.4)**	3.0 (1.7)
	24) …first or last sounds...	**3.7 (0.9)**	**3.7 (1.0)**	**3.4 (1.3)**	**3.2 (1.7)**
	25) …similarities…	2.8 (1.5)	2.6 (1.3)	2.8 (1.3)	2.3 (1.6)
	26) …grouped…	3.0 (1.0)	2.8 (1.2)	3.0 (1.1)	2.7 (1.4)
	27) …English spelling to distinguish...	**3.1 (1.3)**	2.9 (1.1)	2.8 (1.3)	2.3 (1.4)
	28) …Arabic spelling to distinguish…	2.9 (1.4)	2.2 (1.1)	2.0 (1.2)	1.3 (0.6)
Metacognitive	29) …ignored English spelling…	2.6 (1.0)	2.0 (1.0)	2.1 (1.2)	1.4 (0.7)
	30) …ignored Arabic spelling…	2.5 (1.2)	2.4 (1.2)	2.8 (1.6)	2.6 (1.6)
	31) …test my memory...	**3.3 (1.3)**	3.0 (0.9)	**3.5 (1.1)**	**3.3 (1.4)**
	32) …thought about progress...	**3.6 (1.0)**	**3.4 (1.1)**	**3.3 (1.1)**	**3.1 (1.4)**
	33) …noticed mistakes…	**3.5 (1.0)**	**3.7 (0.9)**	**3.6 (0.9)**	**3.3 (1.4)**
Affective	34) …more confident with Arabic spelling…	2.7 (1.2)	2.9 (1.1)	2.8 (1.4)	2.0 (1.4)
	35)…more confident with English spelling…	**3.4 (1.3)**	**3.6 (1.0)**	**3.9 (1.0)**	**4.0 (1.3)**

Turning to the broader cognitive strategies, across all levels there is frequent usage of *remembering the first sound, creating mental images,* and *connecting sounds and letters*, indicating *core* strategies. Meanwhile, all learners rarely *create context* and *break words down into sublexical units*. Figure 6.10 highlights that *creating associations* and *seeking patterns* appears to increase with proficiency, perhaps demonstrating an increased network of associations to draw on across known languages. Additionally, there is a decline in *mouthing* with proficiency, which is logical considering that lower-level learners are at the earlier stages of learning the mechanics of articulating English sounds. Thus, potentially an example of a *base* strategy. The consistent frequency of the metacognitive strategies across proficiency levels, particularly *noticing mistakes,* and *monitoring progress*, may offer evidence that these are *core* strategies (figure 6.11).

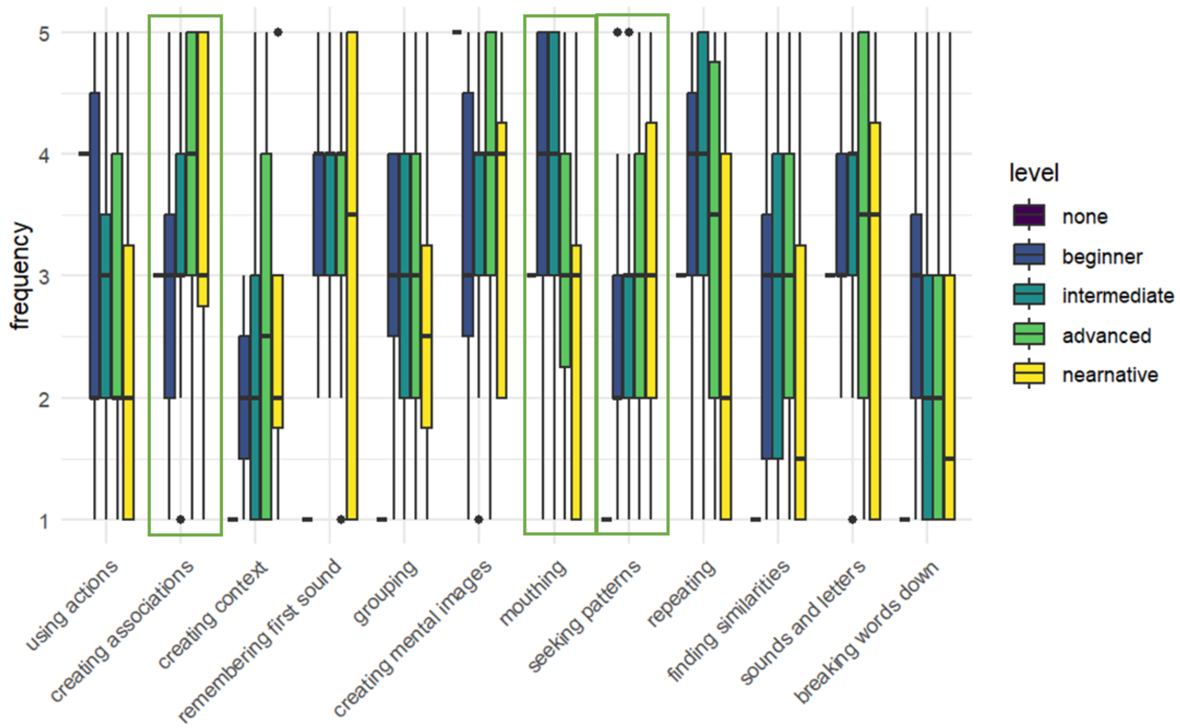


Figure 6.10: Boxplots for frequency of strategy use (scale 1-5) related to cognitive strategies, by proficiency. Frames highlight strategy usage that shifts as proficiency increases.

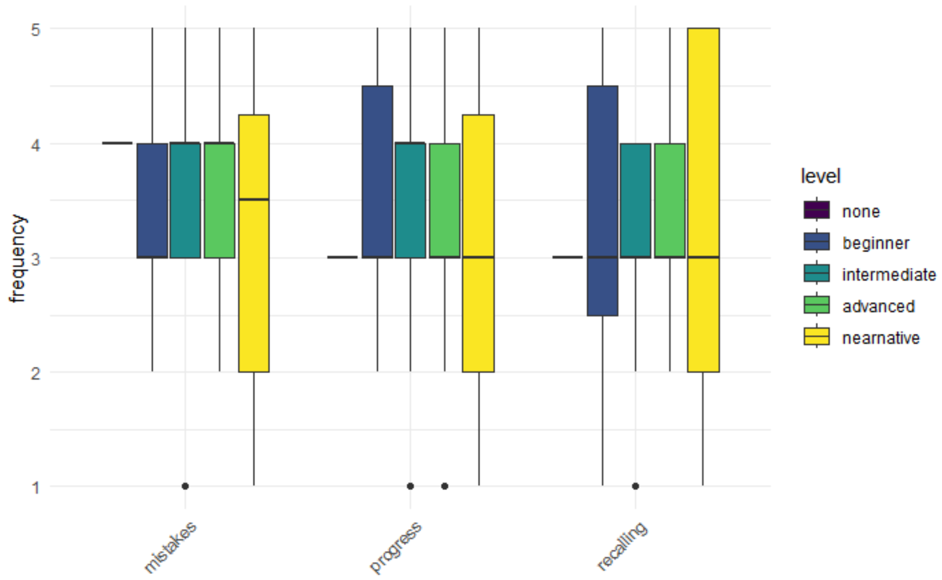


Figure 6.11: Boxplots for frequency of strategy use (scale 1-5) related to metacognitive strategies, by proficiency.

Overall, participants in both language groups report using several different strategies related to phonology and orthography in relation to novel word learning. It is also evident that there is agreement between the strategies identified in the QCA of the open responses (section 6.4.2) and the frequently used strategies reported using the POLLS inventory. For example, participants focused on specific aspects of the input to help remember the word forms, such as the first sounds and letters. Also, participants reported associating, visualising, repeating, and evaluating strategies across both qualitative and quantitative measures. There is also overlap between these strategies and those included in the *Character Learning Strategy Inventory* (Shen, 2005), particularly orthographic-knowledge-based cognitive strategies, which are drawn upon in a modified version of the POLLS inventory, proposed in chapter 8.4. To better understand the dimensions underlying these language learning strategies, responses are further explored with Factor Analysis, specifically using Principal Component Analysis (PCA). The aim of this analysis is twofold: firstly, to locate underlying dimensions (components) of the data, and secondly, to calculate factor scores for participants in relation to the revealed components. These factor scores are then used to explore the relationship between strategy use, lexical encoding accuracy and individual differences, such as proficiency.

6.5.2 PCA with strategy inventory items

There are many similarities, and often much confusion, between PCA and Exploratory Factor Analysis (EFA); however, they are not the same (Plonsky & Gonulal, 2015). PCA is a technique for reducing the dimensionality of data while explaining the maximum amount of total variance of a particular dataset, differing from Factor Analysis which aims to uncover underlying constructs and latent factors to explain the data. The present analysis focuses on investigating the underlying dimensions of reported strategies, in terms of the individual differences within the present sample on this specific task, rather than something more generalisable. Therefore, a PCA approach is sufficient (Dunteman, 1989; Field et al., 2012). As has been the case throughout this study, I assume that the L1 English and Arabic participants are samples from distinct populations, thus PCA is carried out separately with the data of each language group. This analysis is conducted using the principal() function in R. The R scripts for data pre-processing and PCA analysis can be found in appendices XIX and XX, respectively.

6.5.2.1 PCA with L1 Arabic data

The primary focus of this analysis is to understand the components underlying reported LLS, particularly in relation to phonology and orthography when learning novel words. The results for the L1 Arabic group are reported here and later discussed in relation to the L1 English group. Firstly, the correlations between items were assessed, leading to the exclusion of eight items, namely Q14, Q15, Q16, Q18, Q21, Q29, Q30, and Q35[42]. These items either had a low correlation with other items (no correlations of $r > .3$) or the Kaiser-Meyer-Olkin (KMO) result was below .5, indicating sample inadequacy. The remaining 15 items and their correlations are visualised in figure 6.12. The overall KMO result was 0.7 and all values of individual items were >.06, indicating that the sampling was adequate, although "middling" according to Kaiser (1974). Bartlett's test of sphericity indicated that correlations between items were large enough for PCA, χ^2 (105) = 502.3, $p < .001$. Finally, the determinant of the matrix (.007) was sufficiently large to rule out multicollinearity.

[42] These items refer to: creating associations, putting the word in context, creating a mental image, visualising the English spelling, thinking about an action/movement, ignoring English spelling, ignoring Arabic spelling, and feeling more relaxed with English spelling, respectively.

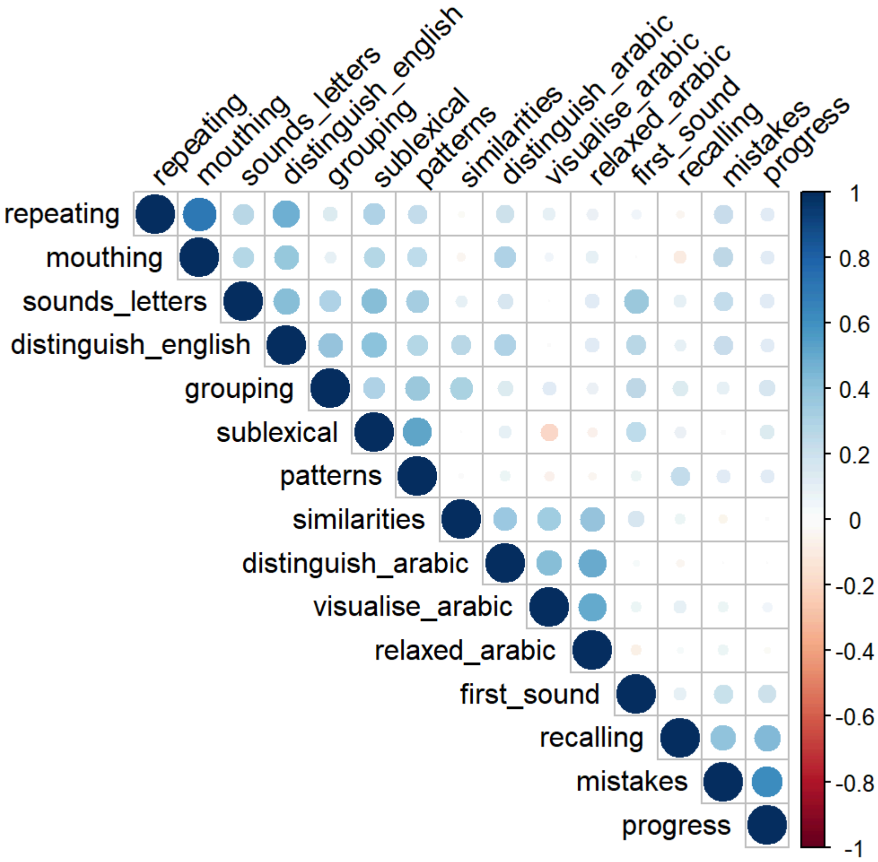


Figure 6.12: Inter-item correlations for the L1 Arabic strategy inventory

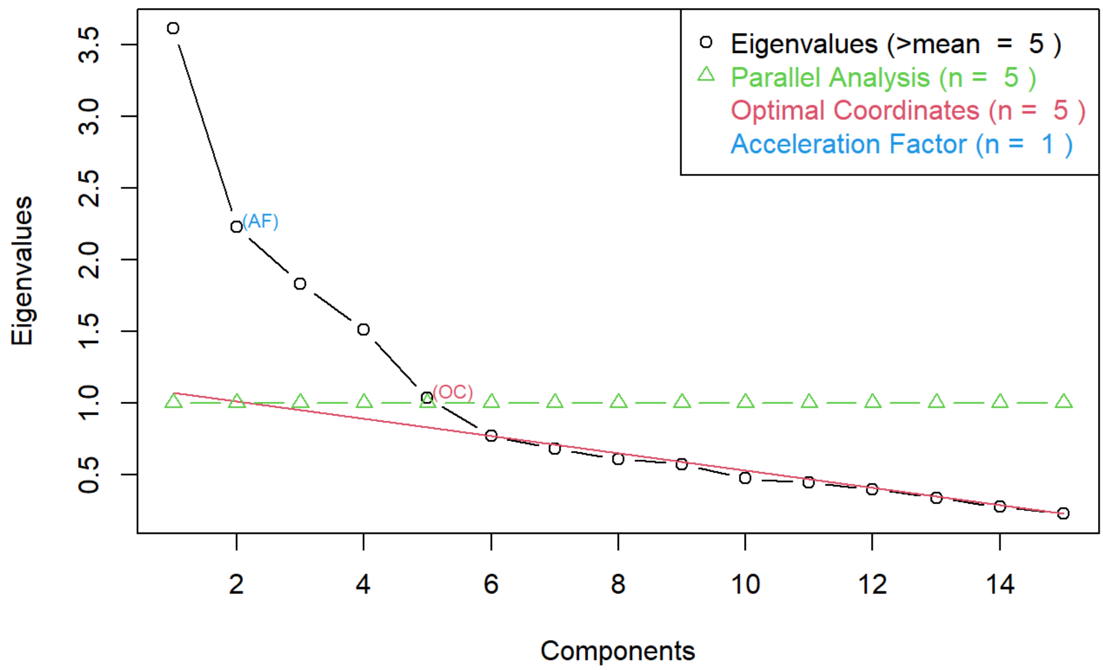


Figure 6.13: Scree plot to assess component extraction for the L1 Arabic data

An initial analysis was run to obtain eigenvalues for each component, which resulted in five components with eigenvalues over Kaiser's criterion of 1, and cumulatively explained 68% of the variance. The scree plot was assessed through the lens of parallel analysis and optimal coordinates, both of which suggested extracting five components in the PCA (Figure 6.13). This decision was further assessed by examining communalities of variables and residuals. When conducting the PCA with five extracted components, the mean communality did not meet Kaiser's criterion of .7 and almost 50% of residuals were greater than 0.05, therefore the analysis was re-run with an additional factor. The subsequent PCA with 6 factors resulted in a mean communality of .73 across all items and a model fit of .93, where the six components accounted for 73% of variance. In addition, the residuals indicated an improved fit, as this time 36% were greater than .05.

To aid interpretation of components, the PCA was conducted with an oblique rotation (oblimin). Based on the arguments made thus far about the relationships between phonology, orthography, and L1 experience, as well as the likely use of multiple strategies in combination, it assumed that underlying dimensions are likely to be related to each other, rather than independent. For this reason, an oblique rotation was chosen over an orthogonal rotation. As oblique rotations assume correlation between variables, table 6.10 reports the factor loadings for both the pattern and structure matrices, which are the regression and correlation coefficients, respectively (Field et al., 2012; Graham et al., 2003). The regression coefficients in the pattern matrix are the unique contributions to a given factor, whereas the structure matrix factor loadings indicate the correlations between factors and variables. In addition, table 6.11 provides the component correlation matrix, indicating some correlation between component 1 and 4 (r = .23), but little relationship between the other components.

While six components were extracted, only four components were indicated by three or more items with a loading > .4, which also do not load highly across multiple other components. This is the recommended minimum requirement for component and factor identification (Child, 2006; Izquierdo et al., 2014). Inadequate sampling of the different domains is a possible explanation here, as there were only a limited number of items in the instrument, which were then further restricted in pre-processing. On this basis and considering the reliability of each subscale, measured using Cronbach's alpha, four components are reported below. The items that clustered on these four components suggested the following dimensions underlying reported language learning strategies:

- Component 1: Producing new words
- Component 2: Using Arabic written input for support
- Component 3: Evaluating learning
- Component 4: Analysing sounds and letters

These four components account for 54% of variance and are further outlined in table 6.12. The script used to run the analysis on the L1 Arabic strategy inventory data in R, including the addition of calculating factor scores, was as follows:

principal(strategy_ArQs2, nfactors = 6, rotate = "oblimin", scores = TRUE)

Table 6.10: L1 Arabic strategy PCA after oblimin rotation. Factor loadings presented for both pattern and structure matrices

	Component				
Item	**1**	**2**	**3**	**4**	**h2**[1]
Q19) I repeated the words…	**.89**[2] **(.89**[3]**)**	.01 (.11)	.06 (.11)	.02 (.22)	.79
Q20) I practised mouth positions...	**.88 (.89)**	.09 (.16)	.04 (.09)	.06 (.24)	.81
Q27) I used English OI to distinguish similar sounds.	**.52 (.59)**	-.03 (.12)	.00 (.11)	.10 (.33)	.68
Q34) I felt more relaxed/confident when I saw the Arabic OI.	-.01 (.08)	**.86 (.85)**	.00 (.01)	.02 (-.02)	.73
Q17) I visualised the Arabic spelling...	-.04 (.01)	**.75 (.77)**	.15 (.14)	-.18 (-.19)	.65
Q28) I used Arabic OI to distinguish similar sounds.	.22 (.32)	**.72 (.75)**	-.13 (-.08)	.09 (.12)	.66
Q32) I thought about my progress...	.08 (.13)	-.06 (-.02)	**.83 (.84)**	-.04 (.09)	.72
Q33) I noticed my mistakes…	.25 (.30)	.01 (.04)	**.81 (.83)**	-.14 (.01)	.79
Q31) I tested my memory of the words.	-.31 (-.19)	.09 (.08)	**.73 (.73)**	.32 (.32)	.71

Q22) I tried to find patterns…	.04 (.22)	-.02 (-.02)	.08 (.17)	**.85 (.87)**	.77
Q21) I broke words down into syllables or sounds.	.14 (.34)	-.10 (-.10)	-.11 (.01)	**.70 (.78)**	.71
Q24) I remembered the first sounds.	-.12 (.03)	-.07 (-.04)	.10 (.21)	-.13 (.02)	.80
Q23) I connected sounds and letters.	.07 (.32)	.18 (.16)	.00 (.14)	**.40 (.49)**	.73
Q26) I grouped similar sounding words.	.05 (.13)	-.09 (.07)	.07 (.16)	.23 (.38)	.71
Q25) I looked for similarities with words and the languages I know.	-.10 (-.07)	.37 **(.50)**	-.08 (-.04)	-.14 (-.06)	.73
Eigenvalues	2.18	2.10	2.00	1.76	
% of variance	14.53	14.00	13.33	11.73	
α	.77	.73	.73	.68	

[1] Communality scores
[2] Regression coefficient from pattern matrix (β)
[3] Correlation coefficient from structure matrix in parentheses (*r*)

Table 6.11: L1 Arabic strategy PCA component correlation matrix

Component	1	2	3	4
1	1.00			
2	0.11	1.00		
3	0.06	0.02	1.00	
4	**0.23**	-0.04	0.11	1.00

Table 6.12: L1 Arabic labelling and interpretation of the variables and components

Comp.	**Variables**	**Label**	**Short Interpretation**
1	Repeating; mouthing; distinguishing sounds with English OI	Producing new words	Learning words through practising the production of sounds - may be guided by English spelling.
2	Confident with Arabic OI; visualising Arabic OI; distinguishing sounds with Arabic OI; looking for similarities with known languages	Using Arabic written input for support	Learning words by drawing on L1 orthographic knowledge and written input, which supports cognitive and affective processes.

3	Thinking about progress; noticing mistakes; testing memory of words	Evaluating learning	Learning words by reflecting on the effectiveness of learning and ways to improve.
4	Seeking patterns; breaking words down into syllables and sounds; connecting sounds and letters	Analysing sound and letters	Learning words through analysing sounds, letters and patterns across the words.

Pattern and structure matrices reveal similar factor loadings. However, the value of looking at both the regression and correlation coefficients with an oblique rotation is highlighted by the moderate correlation of Q25 with component 2, even though its unique contribution to the factor is not above the threshold of .4 in the matrix. This variable, *looking for similarities and contrasts between the words and known languages,* is included in table 6.12, as it is logical that it would correlate with using L1 written input and literacy experience to support learning. The mention of existing orthographic knowledge, as well as exposure to written input, is intentional and draws on observations from the QCA in section 6.4. For example, participants drew on existing orthographic knowledge to visualise words and distinguish between sounds based on speculative spelling, whether or not they had seen the written forms during the learning phase.

Looking at component 1, *producing new words*, it may seem unusual that Q27, *using English OI to distinguish similar sounds*, clusters with the variables related to *repeating* and *mouthing* the words. The factor loadings are lower and analysis of this subscale reveals that the reliability could be improved by removing Q27. However, it was decided to continue with the inclusion of Q27, as this aligned with findings reported in chapter 5.5.2, where L1 Arabic-speakers highlighted the importance of seeing the spelling of a word for correct production. This supports the consistent claim, argued throughout this thesis, that orthography and phonology are closely intertwined in the ways that participants approach word learning.

Component 4 also demonstrates the connection between phonology and orthography in the context of *analysing sounds and letters*. The low but evident correlation between component 1 and component 4 suggests a connection between the visual analysis of written forms and pronunciation for these learners. The only component not to explicitly mention orthographic or phonological knowledge is the metacognitive strategy of evaluating learning. Overall, it appears that learners actively draw on written language experience and exposure when learning the new words, even though there remains a lack of evidence that written input supports learning in the ways that learners perceive or intend it to. To further explore this, participant factor scores were calculated for each of the extracted components and analysed in relation to matching task performance and proficiency, reported in section 6.5.3. First, a similar PCA was conducted with the L1 English LLS data.

6.5.2.2 PCA with L1 English data

To provide exploratory comparison, the same analysis was conducted with the L1 English group responses. It was anticipated that the language and literacy background of this group would result in different dimensions underlying reported LLS, compared to the L1 Arabic group. Specifically, the unintelligibility of the Arabic script input and the fact that all target contrasts are well-established in L1 English phonology are notable distinctions between the two groups. The analysis followed the same procedure as reported with the L1 Arabic data, which began by assessing the correlation between items. This led to the exclusion of two items, namely Q21, Q30.[43] This was because these items either had a low correlation with the other items or due to sampling inadequacy, where the KMO result was below .5. The remaining 21 items and their correlations are visualised in figure 6.14. The overall KMO result was 0.66 and all values of individual items were >.5, which indicated that the sampling was adequate, although "mediocre" according to Kaiser (1974) and lower than the L1 Arabic data. Bartlett's test of sphericity indicated that correlations between items were large enough for PCA, χ^2 (210) = 669.7, $p < .001$. Additionally, the determinant of the matrix (.002) was sufficiently large to rule out multicollinearity.

[43] These items are related to: breaking words down into smaller units and ignoring Arabic spelling.

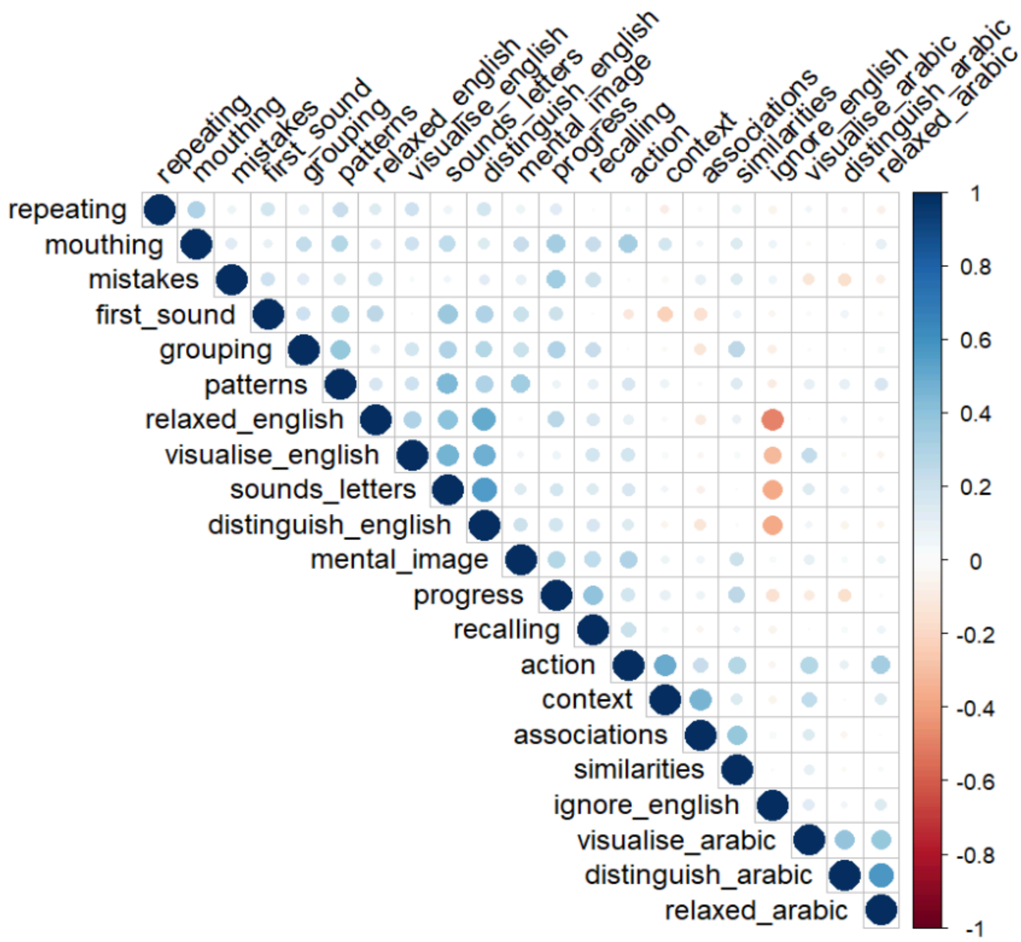


Figure 6.14: Inter-item correlations for the L1 English strategy inventory

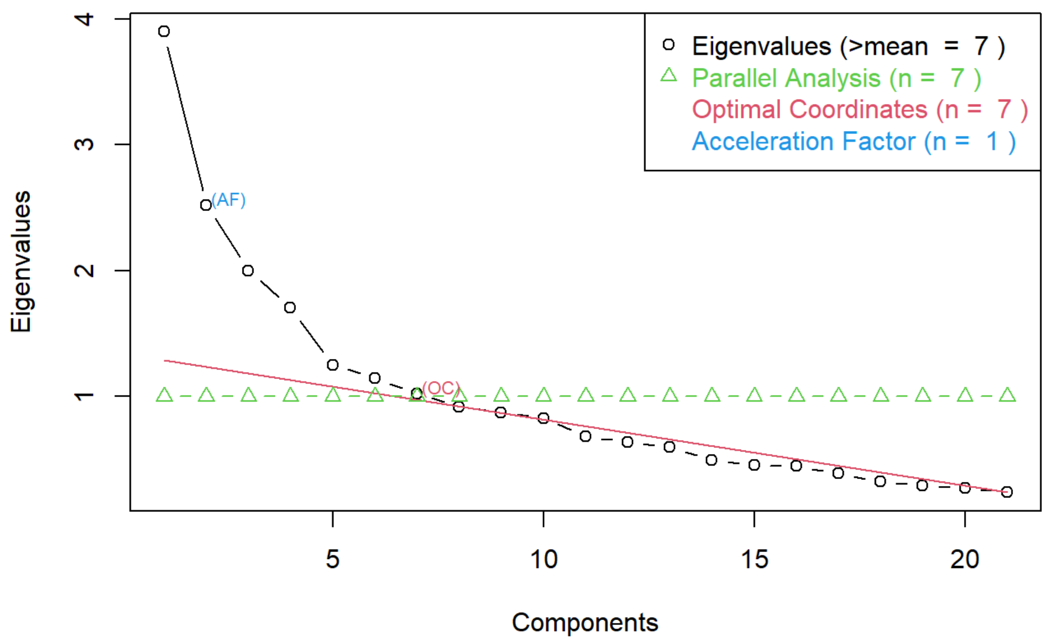


Figure 6.15: Scree plot to assess component extraction for the L1 English data

As before, an initial analysis was run to obtain eigenvalues for each component, resulting in seven components with eigenvalues over 1, cumulatively explaining 65% of variance. The scree plot was assessed with parallel analysis and optimal coordinates, which suggested continuing with PCA extracting seven components (figure 6.15). This was further assessed based on residuals and communalities of variables. When conducting the PCA with seven components extracted, the model fit was below .9, mean communality did not meet Kaiser's criterion of .7 and 44% of residuals were greater than 0.05, therefore the analysis was re-run with an additional extracted component. While the residuals indicated improved fit, the model fit remained below .9 and mean communality below .7, therefore the analysis was conducted again with nine components extracted. This analysis resulted in a mean communality of .73 across all items and a model fit of .91, where the nine components accounted for 73% of variance. In addition, the residuals indicated an improved fit, as this time 33% were greater than .05.

For the same reasons as the L1 Arabic data, the PCA was conducted with an oblique rotation (oblimin). Table 6.13 shows the factor loadings on both the pattern and structure matrices, where structure correlation coefficients are in parentheses. Table 6.14 shows the component correlation matrix, indicating some correlation between component 2 and 3 (r =.2), but almost no relationship between the other components. While nine components were extracted, only four components were indicated by three or more items with a loading > .4, which do not load highly across multiple components. Therefore, only these four components are reported below and are described in more detail in table 6.15. As with the L1 Arabic data, the quality of component extraction is likely due inadequate sampling of the different domains based on a limited number of items. More items were included in this analysis compared to the L1 Arabic analysis. However, more comprehensive inventories, like SILL, consist of 50 items (Oxford, 1990) and an additional 20 were added to the modified SILL (Peterson, 1997). This should be considered regarding interpretation of this PCA and future research applications of the POLLS inventory. The items that clustered on the four components suggest the following dimensions underlying reported LLS:

- Component 1: Connecting words to meaning
- Component 2: Analysing sounds and letters

- Component 3: Using English written input for support
- Component 4: Using Arabic written input for support

These four components account for 39% of variance, which is less than desirable but still of interest comparatively with the L1 Arabic data. The script used to run the analysis on the L1 English strategy data in R, including the additional factor score calculation, was as follows:

principal(strategy_EnQs2, nfactors = 9, rotate = "oblimin", scores = TRUE)

Table 6.13: L1 English strategy PCA after oblimin rotation. Factor loadings presented for both pattern and structure matrices

	Component				
Item	**1**	**2**	**3**	**4**	**h2**[1]
Q15) I put the word in a context…	**.84**[2] **(.83**[3]**)**	.04 (-.02)	.07 (.01)	-.02 (.07)	.71
Q21) I thought about an action…	**.72 (.74)**	.13 (.18)	.05 (.05)	.19 (.28)	.70
Q14) I created associations...	**.59 (.64)**	-.11 (-.16)	.06 (-12.)	.11 (-.06)	.71
Q22) I tried to find patterns…	.07 (.05)	**.74 (.76)**	-.04 (.11)	.04 (.13)	.64
Q16) I made a mental image...	.23 (.22)	**.66 (.63)**	.15 (-.12)	-.18 (.05)	.58
Q24) I remembered the first sounds.	-.27 (-.33)	**.60 (.62)**	.18 (.27)	.10 (.08)	.73
Q23) I connected sounds and letters.	.06 (.02)	**.54 (.63)**	**.40 (.56)**	.00 (.05)	.70
Q29) I ignored the English OI.	-.07 (-.02)	.05 (-.08)	**-.85 (-.82)**	.10 (.16)	.76
Q35) I felt more relaxed/confident when I saw the English OI.	.01 (-.01)	-.01 (.21)	**.80 (.82)**	.16 (.08)	.75
Q27) I used English OI to distinguish similar sounds.	-.04 (-.08)	.39 **(.52)**	**.42 (.60)**	-.11 (-.08)	.68
Q28) I used Arabic OI to distinguish similar sounds.	-.12 (-.03)	-.02 (.07)	.06 (-.03)	**.88 (.86)**	.77
Q34) I felt more relaxed/confident when I saw the Arabic OI.	.13 (.21)	.02 (.10)	.01 (-.08)	**.84 (.85)**	.76
Q31) I tested my memory of the words.	-.01 (.07)	-.05 (.14)	-.02 (.08)	.06 (.05)	.72

Q32) I thought about my progress...	.08 (.12)	.00 (.21)	.26 (.30)	-.04 (-.10)	.70
Q18) I visualised the English spelling...	.00 (.04)	-.01 (.17)	.16 (.35)	-.12 (-.07)	.82
Q17) I visualised the Arabic spelling...	.15 (.27)	.02 (.07)	-.22 (-.17)	**.51 (.60)**	.75
Q19) I repeated the words…	-.12 (-.10)	-.06 (.12)	-.01 (.09)	-.02 (-.04)	.81
Q20) I practised mouth positions...	.27 (.28)	.16 (.32)	-.05 (.04)	.00 (.03)	.72
Q25) I looked for similarities with words and the languages I know.	.12 (.25)	-.04 (.09)	.04 (.02)	.04 (.03)	.80
Q26) I grouped similar sounding words.	-.34 (-.24)	.34 **(.50)**	-.01 (.11)	-.01 (-.01)	.77
Q33) I noticed my mistakes…	-.05 (-.03)	.05 (.14)	-.05 (.05)	-.14 (-.18)	.74
Eigenvalues	2.04	2.14	2.07	1.92	
% of variance	9.71	10.19	9.86	9.14	
α	.65	.63	.76	.67	

[1] Communality scores
[2] Regression coefficient from pattern matrix
[3] Correlation coefficient from structure matrix in parentheses (*r*)

Table 6.14: L1 English strategy PCA component correlation matrix

Component	1	2	3	4
1	1.00			
2	-0.06	1.00		
3	-0.06	**0.20**	1.00	
4	0.11	0.09	-0.09	1.00

Table 6.15: L1 English labelling and interpretation of the variables and components

Comp.	Variables	Label	Short Interpretation
1	Creating context; using actions; creating associations	Connecting words to meaning	Learning words through associations, based on familiar words and/or objects.
2	Finding patterns; making a mental image; remembering first sounds; connecting sounds and letters; distinguishing sounds with English OI; grouping	Analysing sound and letters	Learning words through sublexical analysis and patterns across the words, including grapheme - phoneme correspondences.

3	(not) ignoring English OI; confidence with English OI; distinguishing sounds with English OI; connecting sounds and letters	Using English written input for support	Learning words by drawing on L1 orthographic knowledge and written input, which supports cognitive and affective processes.
4	Distinguishing sounds with Arabic OI; confidence with Arabic OI; visualising Arabic OI	Using Arabic written input for support	Learning words by drawing on L2 orthographic knowledge and written input, which supports cognitive and affective processes.

Pattern and structure matrices reveal similar factor loadings, although it is noteworthy that Q26 and Q27 are moderately correlated with component 2, even though the unique contribution to the factor was not above .4 in the pattern matrix. Both of these variables are included in the variable list in table 6.15, as it is logical that *grouping similar sounding words* and *using English OI to distinguish between similar sounds* would correlate with *analysing sounds and letters*. The same label and short interpretation has been given to this component as component 4 for the L1 Arabic data, due to the overlap between the two components and the common ground of the correlating variables. The focus of the components in both the L1 Arabic and English data appears to centre around looking for patterns between words, sublexical units of sound and the connection between sounds and letters. The addition of *using mental images* in this component may then relate to qualitative findings in chapter 5 and section 6.4, where participants visualised the spelling and image in association with the sound of words. The correlation matrix in table 6.14 highlights that components 2 and 3 are correlated with each other, but not the other factors. This is logical as both involve interpreting familiar sound-symbol correspondences and draw on this knowledge to support word learning.

The most reliable subscale of component 3, *using English written input for support*, mirrored that of component 2 in the L1 Arabic data, which instead referred to Arabic written forms. Taken together, across languages and writing systems, participants drew on L1 literacy to support word learning. It is interesting that the tactics related to L2 English literacy for the L1 Arabic group were not confined to a single strategy component, but integrated with strategies, such as *producing new words*. Meanwhile, the L1 English use of tactics drawing on unfamiliar Arabic OI clustered together to form the fourth component, rather than being integrated with other strategies. Looking at the descriptive data in section 6.5.1, this most likely reflects correlated responses of low engagement with Arabic written forms. The final component to mention is, *connecting words to meaning*, which was less apparent in the L1 Arabic data. The presence of this strategy aligns with discussion around sound-image referential meanings established by learners, reflected in a large proportion of open responses from the L1 English group, discussed in section 6.4. These responses emphasised creating associations and meaningful contexts to learn the new words. However, cautious interpretation of the subscales which tested below $\alpha = .7$ for reliability is recommended, which includes all except component 3 for the L1 English data.

6.5.2.3 PCA summary

The PCA analyses reported above reduced the dimensions of the POLLS strategy inventory down from 23 different variables to four key components for both language groups. Focusing on the results from the L1 Arabic group, the PCA extracted four key strategy components from the data, which were interpretively labelled as (1) *producing new words*, (2) *using Arabic written input for support*, (3) *evaluating learning*, and (4) *analysing sounds and letters*. The subscales of each component were found to be reliable with $\alpha > .7$, except for *analysing sounds and letters* ($\alpha = .68$), demonstrating a good level of reliability.

As with the QCA reported in the previous section, the variables clustering on these components indicate that participants were actively drawing on phonological and orthographic knowledge in their strategy usage. Phonology and orthography appeared to be intertwined across all key components, except for *evaluating learning*. It was also of interest that support from L1 Arabic written forms was extracted as one component, whereas the use of L2 English orthographic knowledge was spread across different strategies, indicating that knowledge of English and Arabic literacy was applied to the task in different ways. Items clustering around the supportiveness of L1 Arabic forms highlighted affective and cognitive aspects of this strategy, such as feeling more confident, visualisation and disambiguation of similar sounds. Each of these aspects reflects ideas consistently highlighted in the open responses of participants, in this and the previous chapter. Further, the connection between

L2 English written forms and repeating or mouthing new words mirrors participants' beliefs about the importance of written input for correct production, discussed in chapter 5.5.

The same analysis was conducted with the L1 English data, resulting in the extraction of four key components from the data, namely (1) *connecting words to meaning*, (2) *analysing sounds and letters*, (3) *using English written input for support*, and (4) *using Arabic written input for support*. The subscales of each component were found to be slightly less reliable with $\alpha > .6$, except for *using English written input for support* which had higher reliability ($\alpha = .76$). Thus, one of the most reliable cross-linguistic strategies was the use of L1 orthographic knowledge when learning novel words. Notably, the affective support of seeing written forms in an L1 shared script and the usefulness for differentiating confusable sounds were common contributing items to this component for both language groups. Both groups also shared similar strategy components of *analysing sounds and letters*, but then differed in the L1 English focus on *connecting words to meaning* and the low usage of strategies related to Arabic written forms.

These results demonstrate that participants consciously use a range of language learning strategies, often involving phonological and orthographic knowledge. Strategy usage differs between the L1 Arabic and L1 English group, but both groups combine phonological and orthographic knowledge within strategies and particularly draw on L1 literacy during novel word learning. In the next section, the reduced dimensions of the L1 Arabic LLS are further analysed in relation to participants' ability to lexically encode the target contrasts in the matching task, as well as individual differences, such as proficiency.

6.5.3 Relationships between strategies, proficiency, and lexical encoding

Based on the data presented thus far, even though written input was not found to facilitate the lexical encoding of a confusable L2 contrast (chapter 4), most L1 Arabic participants considered it important to learn new words accompanied by written input, with a strong preference for learning with English spelling (chapter 5). These beliefs and preferences are reflected in the strategies identified so far in this chapter. It is then of interest to know how the use of different strategies relates to task performance and whether further light can be shed on the apparent discrepancy between participant perceptions about their learning and their actual performance, when learning words with written input. The following sections present correlation and regression analyses conducted with factor scores, matching task accuracy and learner characteristics. Factor scores for each participant were calculated for the key components extracted from the L1 Arabic PCA in section 6.5.2.1, and then analysed to investigate: (1) correlations amongst the variables of interest, (2) whether specific strategies predict performance on the matching task, and (3) whether learner characteristics predict strategy usage. Throughout, there is a focus on L2 English proficiency as this was a significant predictor during the initial analysis in chapter 4. The R script for the correlation and regression analyses can be found in appendix XXI .

6.5.3.1 Correlations analysis of individual differences and accuracy

The data was initially explored using correlation analysis, to investigate relationships between strategy usage, learner characteristics and task performance. Specifically, the four extracted strategies components (*producing new words, using Arabic written input for support, evaluating learning* and *analysing sounds and letters*) were assessed in relation to d-prime scores[44] from the matching task. Additionally, variables of proficiency, exposure to English literacy, amount of distraction and age were included, based on their relevance to analyses reported in previous chapters. Figure 6.16 visualises significant correlation coefficients, using the Spearman method for nonparametric data.

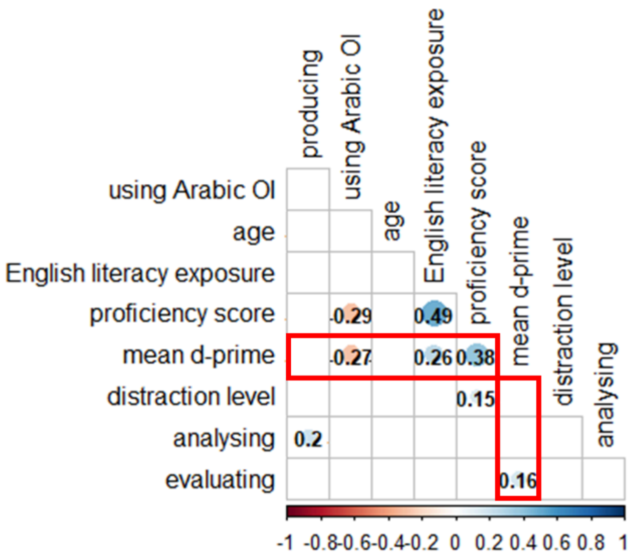


Figure 6.16: Correlations between learner characteristics, strategies, and accuracy

[44] d-prime scores were chosen as a better measure of sensitivity to the target phonological contrasts, compared to raw correct scores.

The red boxes highlight the variables that significantly correlate with matching task accuracy (mean *d'*). As anticipated, English proficiency test scores are positively correlated with accuracy, r =.38, p <.001, as is daily exposure to English literacy, r = .26, p =.006. English literacy exposure and proficiency are also significantly correlated with each other, r =.49, p <.001, which logically implies that higher proficiency learners engage in more regular L2 English literacy practices than lower proficiency learners. Accuracy and *evaluating learning* are then positively, but not significantly, correlated, r =.16, p = .09. Meanwhile, accuracy and *using Arabic written input for support* are negatively correlated, r = -.27, p = .004. This strategy component is also negatively correlated with proficiency, r =-.29, p=.002. The other strategy components, *producing new words* and *analysing sounds and letters*, are only correlated with each other, r =.20, p =.04, which was also highlighted in the PCA analysis in section 6.5.2.1. The relationship between LLS components and *d'* scores is visualised in figure 6.17. This analysis provides preliminary evidence that the usage of different LLS is not strongly related to task performance, apart from some indication that reliance on L1 Arabic orthography is correlated with worse performance and metacognitive evaluation strategies may correlate with improved performance. However, the latter is not well-supported by the data.

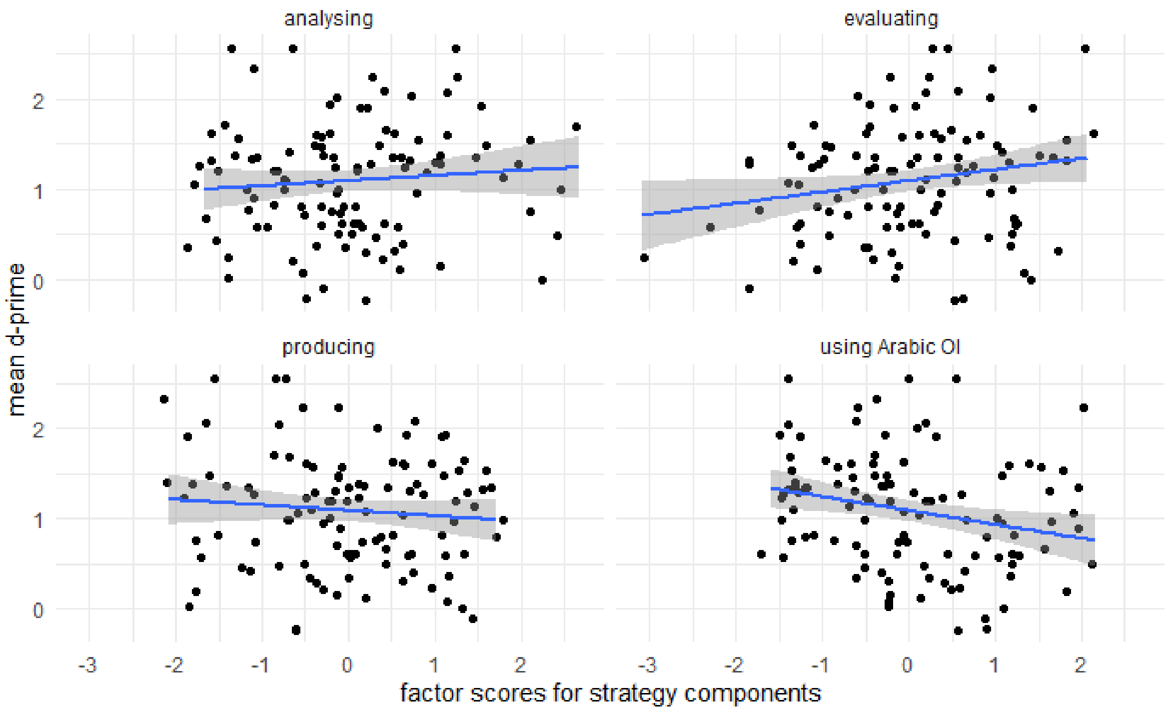


Figure 6.17: Correlations between factor scores for the LLS components and mean *d'* scores in the matching task

6.5.3.2 Multiple regression analysis of accuracy and strategies

Taking this analysis a step further, multiple regression analysis was conducted to ascertain whether strategy component factor scores predicted task accuracy. L2 English proficiency was also included in the model, based on the significant correlation with accuracy reported in section 6.5.3.1. As multiple regression assumes no multicollinearity, exposure to English literacy was not included in this model, based on the high correlation with proficiency. Age and proportion of distraction were included in a preliminary model, but were not significant predictors of accuracy and did not improve the model fit. A hierarchical approach was adopted, where the known predictor of L2 proficiency was entered first, followed by the factor scores for the strategy components in the second step. Table 6.16 reports the results and the script used to run the analysis on the L1 Arabic data in R was as follows:

```
model1 <- lm ( meanD.prime ~ prof_test_score, data = strategy_scoresAr2 )
model2 <- lm ( meanD.prime ~ prof_test_score +
                    producing + using ArabicOI + evaluating + analysing,
                    data = strategy_scoresAr2 )
```

Table 6.16: Multiple hierarchical regression predicting lexical encoding accuracy, as measured by participant mean *d'* scores

		Adj. R^2	**Estimate**	**Std. Error**	**β**	***p***
Step 1		0.13				
	(Intercept)		0.04	0.26		.866
	Proficiency		0.12	0.03	0.37	<.001***
Step 2		0.17				
	(Intercept)		0.18	0.27		.493
	Proficiency		0.09	0.03	0.32	<.001***
	Producing new words		-0.03	0.06	-0.05	.613
	Using Arabic OI		-0.09	0.06	-0.15	.108
	Evaluating learning		0.13	0.06	0.21	.02*
	Analysing sounds & letters		0.02	0.06	0.03	.707

The results in table 6.16 show that proficiency is indeed a significant predictor of participant accuracy, which alone accounted for 13% of variance in the data. By adding the strategy components to the model, the adjusted R^2 only increases by 4%, explaining 17% of total variance. *Evaluating learning* is also a significant predictor of lexical encoding accuracy, where higher frequency usage of evaluating strategies predicts improved signal detection sensitivity (*d'*). However, model comparison using the *anova()* function shows the inclusion of the strategy components does not significantly improve model fit ($\chi^2(4)=2.13$, $p=0.08$). Outlier and assumption checks for this analysis are provided in appendix XXI.

According to these results, there is no evidence that the use of strategies involving orthographic or phonological knowledge predicts better lexical encoding of confusable L2 contrasts. Indeed, there is little evidence that strategy usage, in general, has any significant influence on word learning accuracy, echoing the findings of Bassetti and colleagues (2020). The limitations of the POLLS inventory have already been stated, particularly in terms of the inventory size. Thus, this analysis would benefit from replication with a more robust testing instrument. Additionally, it would be of interest to explore more longitudinal effects, to see how strategy usage impacts word learning over a longer period, and with a design less centred on short-term memory.

6.5.3.3 Linear regressions of strategies by proficiency

Having established that strategy use does not predict improved performance, it is of interest whether learner characteristics can shed more light on participant strategy usage. Looking at the correlations in section 6.5.3.1, the only notable relationship was between proficiency and the strategy component *using Arabic written input to support learning*. Linear regression analyses were then conducted to further investigate whether L2 English proficiency predicts the usage of the extracted strategy components. Table 6.17 reports the results and the script used to run the analysis in R was as follows:

```
model_producing <- lm (producing ~ prof_test_score, data = strategy_scoresAr2)
model_usingArabicOI <- lm (usingArabicOI ~ prof_test_score, data = strategy_scoresAr2)
model_evaluating <- lm (evaluating ~ prof_test_score, data = strategy_scoresAr2)
model_analysing <- lm (analysing ~ prof_test_score, data = strategy_scoresAr2)
```

Table 6.17: Linear regressions of the relationship between L2 English proficiency and strategy usage, based on PCA factor scores

	R^2	Estimate	Std. Error	β	*p*
model_producing	0.03				
(Intercept)		0.73	0.43		.093
Proficiency		-0.08	0.04	-0.17	.085
model_usingArabicOI	0.08				
(Intercept)		1.28	0.42		.003**
Proficiency		-0.13	0.04	-0.29	.002**
model_evaluating	0.00				
(Intercept)		0.10	0.44		.828
Proficiency		-0.01	0.04	-0.02	.824
model_analysing	0.01				
(Intercept)		-0.52	0.44		.234
Proficiency		0.05	0.04	0.12	.222

The only strategy component that is significantly predicted by L2 English proficiency is *using Arabic written input for support*, where use of this strategy declines as L2 English proficiency increases. This clarifies the previously mentioned negative correlation with *d′* scores and demonstrates that lower proficiency learners are more likely to rely on L1 orthography. It is then their English proficiency that predicts both the adoption of these strategies and their lower accuracy on the matching task. This may also reflect gradually automatised application of L2 over L1 orthographic knowledge with increasing proficiency, as proposed by Shen's (2005) research with the *Character Learning Strategy Inventory*. Proficiency did not predict usage of any other strategies, aligning with discussion of descriptive results in 6.5.1, where *using Arabic written input for support* may be an example of a *base* strategy, whereas the others are *core* strategies, used across the board. These preliminary findings outline opportunities for further research into L2 orthography and phonology in the context of LLS and individual differences.

6.6 Concluding remarks

This chapter opened with the question: What is the relationship between learners' reported LLS and lexical encoding of L2 phonological contrasts in novel words? Considering the research relevant to orthography and phonology in the context of LLS, including findings reported in chapter 4 and 5, the following predictions were made:

1. Participants will report using strategies related to both written and spoken input for the purpose of L2 phonology and word learning.
2. Cognitive and metacognitive strategies will be more commonly reported in the context of the language learning task in this study.
3. Strategy usage will predict matching task accuracy and relate to L2 proficiency.

The focus of the POLLS inventory on phonological and orthographic knowledge in relation to strategy use offered insight into the number of related strategies and frequency of use, as well as the ways phonology and orthography were often intertwined. For example, descriptive frequencies provided further evidence of L1 Arabic-speaker preferences for learning the target items with English written forms, even more so than the L1 English group. This is a consistent finding across both qualitative and quantitative results, in this and the previous chapter. Additionally, overviews of both qualitative and quantitative data emphasise that the conscious use of written input and orthographic knowledge in LLS usage is the norm, irrespective of proficiency level or language background.

Reducing the dimensions of the POLLS inventory, PCA key components demonstrated how strategies drawing on orthographic and phonological knowledge were often found together, within a single strategy component. Qualitative data illustrated that participants often conflated sounds and letters, as well as intentionally using orthographic knowledge to analyse sounds and create associations with meaning. Open responses also clarified that orthographic influence was not limited to written input exposure, but included the use of broader orthographic knowledge to filter sounds and remember the target items. The examples of speculative visualisation of spellings, whether or not written input was presented, revealed how knowledge of grapheme-phoneme correspondences contributed to lexical representations.

The qualitative and quantitative analysis of participants' self-reported strategy usage revealed that both language groups used a range of LLS. Identified strategies were predominantly cognitive, aligning with previous LLS research. Strategies related to *associating*, *remembering*, *repeating*, *visualising* and *comparing/contrasting* were identified in the QCA of open responses from both language groups. These results overlapped with PCA results and the components of *producing new words* and *analysing sounds and letters* for the L1 Arabic data, and then *connecting words to meaning* and *analysing sounds and letters* for the L1 English data. Additionally, the PCA revealed a specific strategy component connected to L1 orthographic knowledge supporting word learning for both language groups.

Metacognitive strategies were also identified in qualitative and quantitative analyses, particularly in relation to evaluating learning. Additionally, the emphasis on *focusing* strategies within the QCA of open responses demonstrated the conscious attentional resources that were applied to learning the phonological forms of the words. Participant reflections on how they directed their attention also provided insight into points raised in chapter 5, regarding cognitive overload and navigating the potential for written forms to distract from other aspects of the input. Only two items in the POLLS inventory made reference to paying attention, namely "I purposefully ignored the English/Arabic spelling", which should be expanded in future adaptations of this instrument. No other domains were evident based on the qualitative or quantitative analysis, except responses to the items "I feel more confident/relaxed when I see the spelling in…" revealed that participants made use of written input as an affective strategy, or were at least aware of an affective influence. Again, as only two items related to affect, little more can be said on this topic, apart from the fact it deserves further investigation.

The identification of these LLS was conducted with the view to understand whether strategy use influenced L1 Arabic participant performance on the matching task, and whether learner characteristics influenced which strategies were used. Multiple regression analysis confirmed that L2 proficiency significantly predicted lexical encoding accuracy, as also reported in chapter 4. Adding the strategy components to the model showed *evaluating learning* to be a significant predictor of performance accuracy, as well. However, the inclusion of this variable did not significantly improve the model fit. Therefore, there is no evidence that strategy use influenced performance on the experimental task. These results further demonstrate that participants' perceptions of their learning and the strategies they employ involving written forms do not match up with their ability to lexically encode confusable L2 contrasts with orthographic input. Linear regressions revealed that proficiency was also a predictor of strategy usage. However, this was only evident with *using Arabic OI to support learning*, where usage declined as proficiency improved. While use of this strategy was also negatively correlated with matching task accuracy, the regression analyses clarified that this relationship was better explained by the proficiency of participants, rather than their strategy use.

These preliminary insights into the relationship between language learning strategies, individual differences, and word learning across writing systems deserve more rigorous investigation. With this in mind, further refinement of the POLLS inventory is discussed in chapter 8, incorporating qualitative findings and reflections from conducting the quantitative analysis. Overall, the results presented here and in the previous two chapters make a clear case for the closely intertwined relationship of orthography and phonology in the minds of language learners, where linguistic and literacy experience is drawn upon across writing systems. Participants have clear beliefs and preferences around exposure to written input during language learning, which are reflected in the strategies applied to the task of word learning, but do not appear to be effective with regards to the target outcome. Throughout each of the analysis chapters, L2 English proficiency has consistently offered further illumination to results and the discussion of findings. The next chapter takes a closer look at the role of proficiency in the present study, and other individual differences that merit further consideration.

Chapter 7: Beyond proficiency

7.1 Introduction

A recurrent theme throughout the previous analysis chapters has been the influence of L2 English proficiency, particularly in relation to task performance and strategy usage. For instance, proficiency proved to be a strong predictor of accuracy in the matching task, which increased with higher proficiency. Exploration of the interaction between orthographic input (OI) condition and phonological contrast also revealed that accuracy steadily improved in each OI condition, and for both phonological contrasts, with increasing proficiency levels. However, even the highest proficiency learners struggled to perform above chance with /f-v/ contrast words presented with either English or Arabic written input. It was hypothesised in chapter 4 that this steady improvement could reflect the fact that higher proficiency learners (1) are likely to have more robust phonological representations of the target L2 contrasts, (2) are better able to rapidly integrate the visual information available in the English script through higher L2 literacy, and (3) are better able to assess and suppress attention to Arabic written forms. Despite these proficiency-related advantages, it is of critical interest that even high proficiency participants do not appear to benefit from exposure to written input when lexically encoding confusable contrasts. Chapter 6 then revealed that L2 English proficiency also predicted the use of language learning strategies (LLS) relying on L1 Arabic written forms, where higher proficiency learners were less reliant on these strategies than low proficiency learners. However, strategy usage did not predict encoding accuracy.

In contrast to the other analysis chapters, this present focus on proficiency does not directly address a research question laid out at the start of this thesis. Instead, the analysis and discussion presented here sheds light on the role of proficiency in the current study, and what that means for the interpretation of findings. Through exploration of proficiency, what can be understood about the influence of cross-scriptal orthographic information at different points in learners' developmental trajectory? Additionally, what is it about 'proficiency' that influences task performance and strategy usage, and how does this relate to other individual differences? To address these points of enquiry, a cluster analysis is presented to better understand the dimensions of proficiency within the present study. The resulting groups of learners then direct the choice of learner profiles for a more in-depth look at individual differences (IDs) in relation to participant performance, perceptions and strategies. The emerging discussion makes a strong case for future research to explore the role of proficiency in the context of orthographic influence and L2 phonology. Additionally, it demonstrates the value of moving beyond a single variable to investigate the IDs which contribute to our understanding of proficiency.

7. 2 Research context

Proficiency has received little attention in the field of orthographic influence and L2 phonology (Hayes-Harb & Barrios, 2021), as many studies focus on first exposure and use naïve participants. However, the findings reported here align with emerging work investigating proficiency in relation to orthographic influence across writing systems. For example, Hao and Yang (2021) found that the proficiency of L1 English learners of L2 Mandarin positively predicted performance in an audio-visual matching task, testing segmental and tonal encoding. A similar finding was reported for L1 Cantonese learners of L2 Mandarin tones with perception and production of familiar, real words (Mok et al., 2018). Zhang and Roberts (2021) also found that L2 proficiency predicted performance in Hanzi reading and writing, for both L1 English and L1 Arabic learners of L2 Mandarin.

Further insight is gained through closer inspection of the findings of these studies. Hao and Yang (2021) found advanced learners demonstrated greater sensitivity to tonal contrasts when they were exposed to characters rather than Pinyin spelling. Meanwhile, the opposite was observed with the naïve participants. These findings differ from the present study in that there is an interaction between OI condition and proficiency; however, this reflects a different research focus. Critically, the authors interpreted findings as evidence that novel graphemes facilitate high proficiency L2 tonal encoding more than familiar graphemes. I would argue that these are not novel or unfamiliar graphemes to the advanced learners, as the authors clearly state that “characters were known to the two learner groups…to focus on learning the phonological form and lexical meaning of the target words rather than having to learn new graphemes at the same time” (Hao & Yang, 2021, p. 893). It is for this reason that the terms *shared* and *distinct* orthographic input, rather than *unfamiliar*, have been used in the present study, in relation to experience with L2 literacy experience across writing systems. Overall, I find the authors' hypothesis in the discussion more convincing, where they suggest this pattern reflects learners' increased use of characters, as opposed to pinyin, inside and outside of the classroom as proficiency increases.

Mok et al. (2018) also found an interaction between performance and orthographic influence, by comparing how exposure to Pinyin and Characters affected the perception and production of L2 tones. They found that orthographic influence differed by high and low task performance, which was interpreted to reflect L2 proficiency. In particular, lower proficiency learners were more affected by orthographic information than higher proficiency learners. Zhang and Roberts (2021) also reported an interaction between orthography and L2 proficiency, but this time in relation to script directionality. This effect was only present in phonetic radical application skills, and not Hanzi reading or writing, leading the authors to speculate that this orthographic awareness does not progress in parallel with L2 proficiency, but develops later. Zhang and Roberts also drew attention to the biscriptal literacy of the Arabic speakers, who were literate in both the Arabic and Roman alphabet, in comparison to the English speakers, which is relevant but a often neglected consideration (Vaid, 2022).

Proficiency has also been found to predict orthographic influence on word recognition when the L1 and L2 share the same writing system. Escudero et al. (2014) found that the proficiency of L1 Spanish learners of L2 Dutch predicted word recognition accuracy for congruent but not incongruent written input trials. They drew the conclusion that shared and congruent L1-L2 grapheme-phoneme correspondences (GPCs) were reinforced as proficiency increases. However, L1 GPCs also "lead to persistent interference when they cannot be mapped to the target language, regardless of proficiency" (Escudero et al., 2014, p. 394). Taking a processing perspective, Veivo and Jarvikivi (2013) argued that facilitative effects of orthography are only evident for more advanced learners with established mental representations, based on L1 Finnish spoken word recognition in L2 French. A later study by Veivo and colleagues (2018) extended this line of enquiry using eye-tracking and reported more prominent orthographic bias in lower proficiency learners, which decreased with proficiency. This bias was said to reflect orthographic representations being more robust than phonological representations, and thus activated more during L2 spoken word processing. They then speculated that higher proficiency learners were more able to suppress irrelevant or misleading between-language information. Research investigating orthography-induced transfer of L1 phonological processes during a sentence-reading task also found that L1 processes decreased with higher proficiency (Silveira, 2012).

Taken together, these studies make a strong case that proficiency is an important variable to consider when investigating orthographic influence on phonological acquisition and processing. There is agreement that lower proficiency learners rely more heavily on orthographic input, which shares the script of the L1 writing system, as well as being less able to suppress attention to and activation of interfering L1 orthographic knowledge. In contrast, higher proficiency learners have more cognitive control over linguistic knowledge and processing, which means they are better able to take advantage of the facilitative effects of L2 written input, in both L1 shared and distinct scripts. This is related to more robust L2 phonological representations for higher than lower-level learners, as well as more developed L2 literacy skills. However, there is also evidence that while interference effects from orthography-induced L1 transfer may reduce with increased L2 proficiency, they do not disappear. The present study is well-situated within this emerging research and offers evidence to support preliminary understandings of the relationship between proficiency and orthographic influence. The next sections present proficiency groupings formed using cluster analysis and learner profiles to add context and insight into individual differences.

7.3 Cluster Analysis of proficiency groupings

Several measures of proficiency were included in the present study, mostly self-reported alongside a short proficiency test. Based on the correlations between proficiency measures and matching task accuracy, it was decided to use the proficiency test score as the key proficiency measure within analyses. This was a short test and a rudimentary measure, therefore the construct of proficiency would benefit from closer examination. In particular, it is of interest to know how learners could be grouped by proficiency. Cluster analysis is a multivariate exploratory approach which is useful for grouping data in a way that minimises variation within newly established categories, and maximises variation between categories (Staples & Biber, 2015). Its increasing usage in SLA research led to a recent review of current practice by Crowther et al. (2021), which proposed standards for both conducting and reporting Cluster Analysis. These guidelines were closely followed in the analysis reported here and the relevant R script can be found in appendix XXII.

7.3.1 Analytical approach

The two most popular approaches to cluster analysis used in L2 research are hierarchical and k-means (Staples & Biber, 2015). Within hierarchical cluster analysis (HCA), either an *agglomerative* or *divisive* approach can be employed, with agglomerative methods being more frequent. This approach merges cases based on similarity and distance, where the optimal number of clusters are typically determined after merging. In contrast, K-means clustering requires a particular number of clusters to be requested prior to analysis. This can be based on theoretical motivations or through comparing subsequent solutions with different numbers of clusters. Some also use these two approaches in combination (Crowther et al., 2021). The analysis reported here adopts an HCA approach, as there is no clear theoretical basis for a priori cluster determination. Following the guidance of Crowther et al. (2021), Ward's method was chosen as the linkage method and the R script automatically selected the appropriate distance measure, namely squared Euclidean distance.

The objects of clustering were the L1 Arabic participants and the predictor variables were four measures of L2 English proficiency, namely (1) proficiency test score, (2) self-reported level, (3) self-reported English skills ability, and (4) approximate daily exposure to English. Self-reported level was recoded from a 5-level categorical variable (none, beginner, intermediate, advanced, near-native) to a numeric, ordinal variable[45]. Self-reported English skills ability was the mean score given across English speaking, listening, reading, and writing responses (0-5). Then, daily exposure to English was the proportion of estimated total hours of exposure to any language that a participant reported being exposed to English. As these variables all differ in length, they were scaled using standardised z-scores before further analysis. Regarding recommended sample size, the number of participants (n = 109)[46] was sufficient according to Dolnicar's proposal (2002) for a preferred sample size of $5*2^k$, where ^{k} is number of predictor variables (e.g. $5*2^4 = 80$).[47]

Next, the optimal number of clusters was assessed based on visual analysis using the elbow method, average silhouette method, and gap statistics. Figure 7.1 displays these indices, which indicate that a two-cluster solution may be optimal. Scatterplots for two and

[45] The inclusion of ordinal data is an area of ongoing discussion (Jacques & Biernacki, 2018).
[46] The data includes PCA strategy factor scores, thus the sample size reflects the dataset with missing responses removed.
[47] If the sample were much larger, then k-means would be a more appropriate choice

three cluster solutions (figure 7.2) then demonstrated that the first cluster remains constant, whereas the second cluster is divisible into another large and small cluster. These clusters are clearly distinguishable from each other, but the size of the third cluster suggests that a two-cluster solution is preferable, in line with the other measures reported.

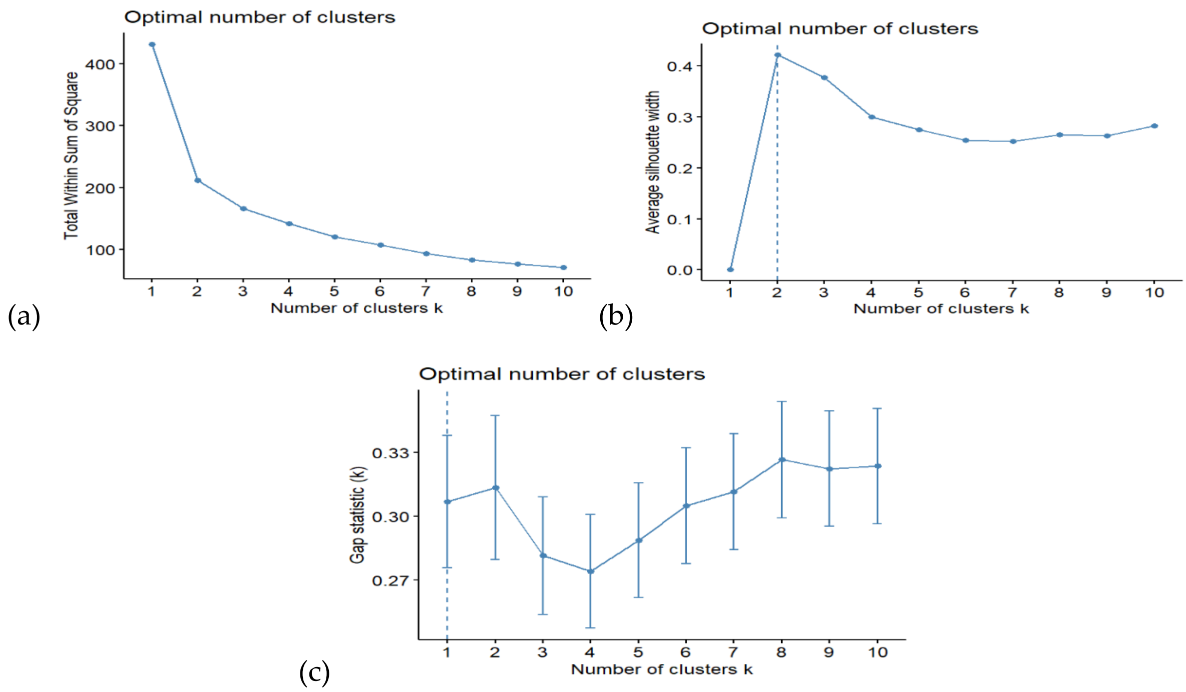


Figure 7.1: Indices to guide optimal cluster decision, namely (a) the elbow method, (b) silhouette method, and (c) gap statistic

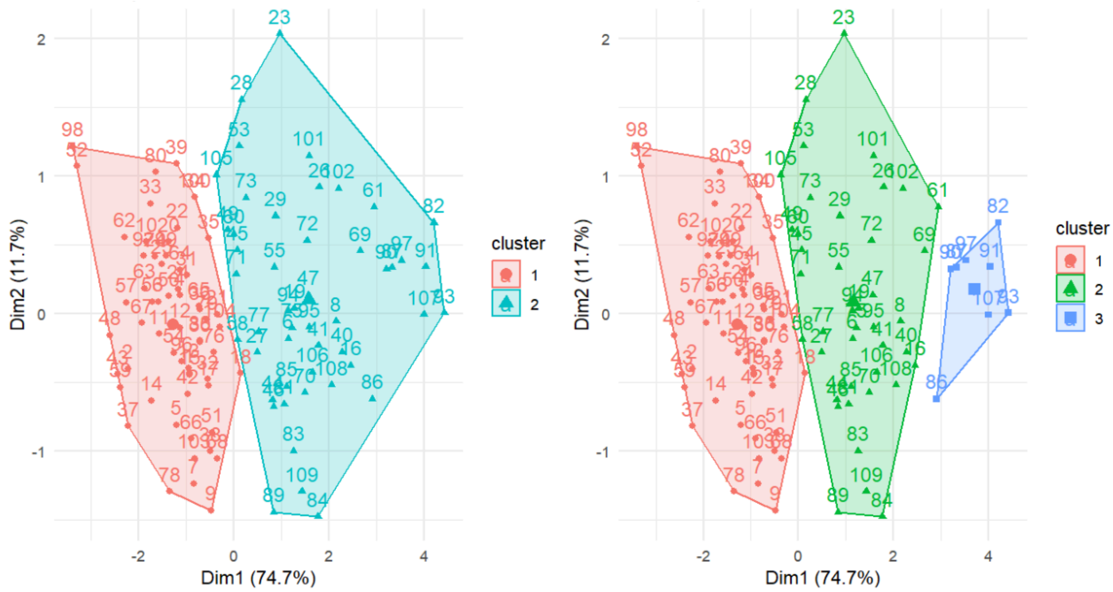


Figure 7.2: Scatterplots for 2 and 3-cluster solutions

7.3.2 Results

Non-parametric Mann-Whitney U tests revealed that the two clusters significantly differed in their proficiency test score, W = 2637, p < .001; self-reported mean skills ability, W = 2790, p < .001; self-reported level, W = 2892, p < .001; and daily exposure to English, W = 2447, p < .001. Figure 7.3 visualises these comparisons by the original scale length, which means that scores are more visually compressed for the variables of level and proportion of daily English exposure. Across the four variables, it is clear that cluster 1 consisted of higher proficiency learners (n=60), whereas cluster 2 consisted of lower proficiency learners (n=45). The higher proficiency group performed more accurately (total correct, W = 1864, p = .02, *d'*, W = 1874, p = .01) with faster response times (RT) in the matching task, compared to the lower proficiency group, although RT difference was not significant. As in chapter 6.5.3.3, application of the strategy *using Arabic OI for support* was higher for the lower proficiency group, W = 1141, p = .05. A summary of descriptive cluster comparisons are presented below for matching task performance (table 7.1) and strategy components factor scores (table 7.2).

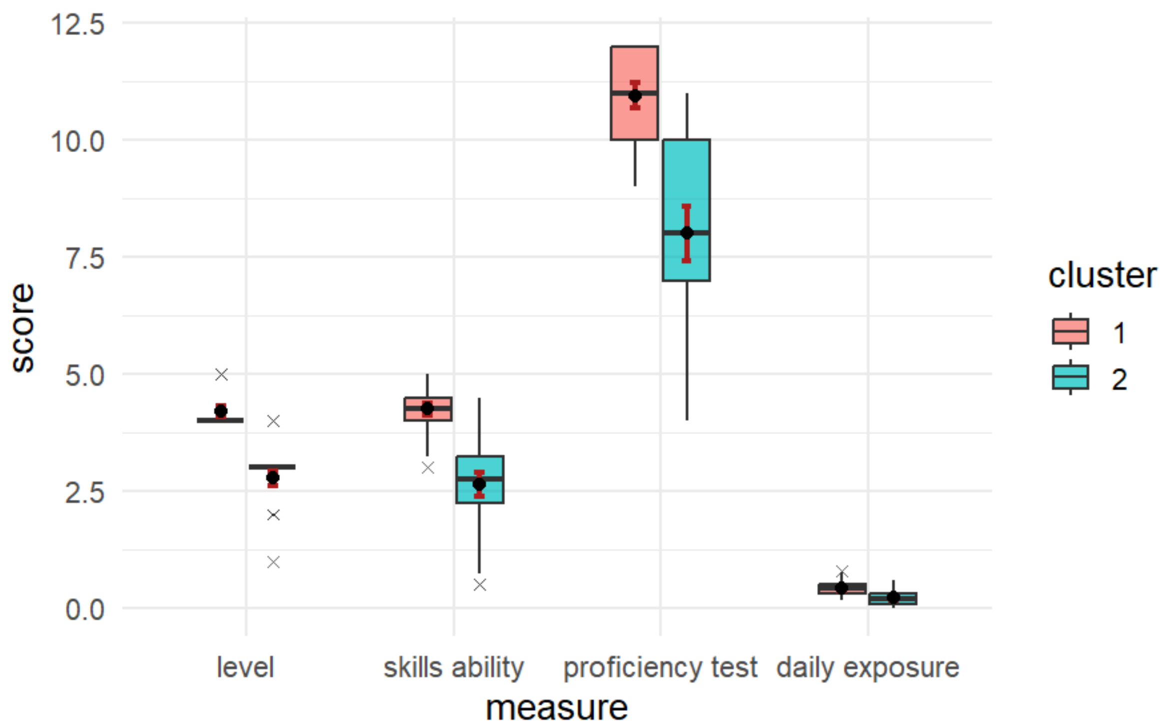


Figure 7.3: Interquartile range, mean and 95% CIs of proficiency scores, by cluster

Table 7.1: Summary of cluster comparisons by proficiency test score, matching task accuracy and response time (RT)

Cluster	Proficiency test score *M* (SD)	% correct *M* (SD)	*d′* *M* (SD)	RT *M* (SD)
1 (*n* = 60)	11 (1.1)	33.5 (5.9)	1.22 (.59)	2093 (848)
2 (*n* = 49)	8 (2.1)	30.9 (6.6)	0.93 (.66)	2286 (877)

Table 7.2: Summary of cluster comparisons by factor scores of LLS components

Cluster	Producing *M* (SD)	Using Arabic OI *M* (SD)	Evaluating *M* (SD)	Analysing *M* (SD)
1 (*n* = 60)	-.16 (1.1)	-.15 (1.1)	.02 (1.0)	.13 (1.0)
2 (*n* = 49)	.20 (0.9)	.18 (0.8)	-.02 (1.0)	-.16 (.9)

These results confirm the patterns related to proficiency reported in chapter 4 and 6, in relation to task performance and strategy use. As proficiency was not expressly included in the central research question, its influence was not consistently explored in relation to all reported findings. In particular, there was no discussion of the relationship between perceived OI influence and beliefs about the importance of exposure to written input in relation to proficiency. Through clustering the data into these proficiency groups, another perspective is made available to revisit this data. Figure 7.4 and 7.5 visualise the proportions of responses to the multiple-choice questions in the posttest questionnaire, by proficiency cluster. The importance of early exposure to OI when learning new words is a strongly held belief, irrespective of proficiency. Similarly, for both higher and lower clusters, there is strong preference for exposure to English spelling when learning the new words. When asked more specifically about participants' experience of Arabic spelling, English spelling and audio-only conditions during the experiment, the only notable difference supports the finding that proficiency influences how participants interact with L1 Arabic written input. Specifically, lower proficiency participants more frequently reported that Arabic OI made learning the words easier, compared to the higher proficiency group.

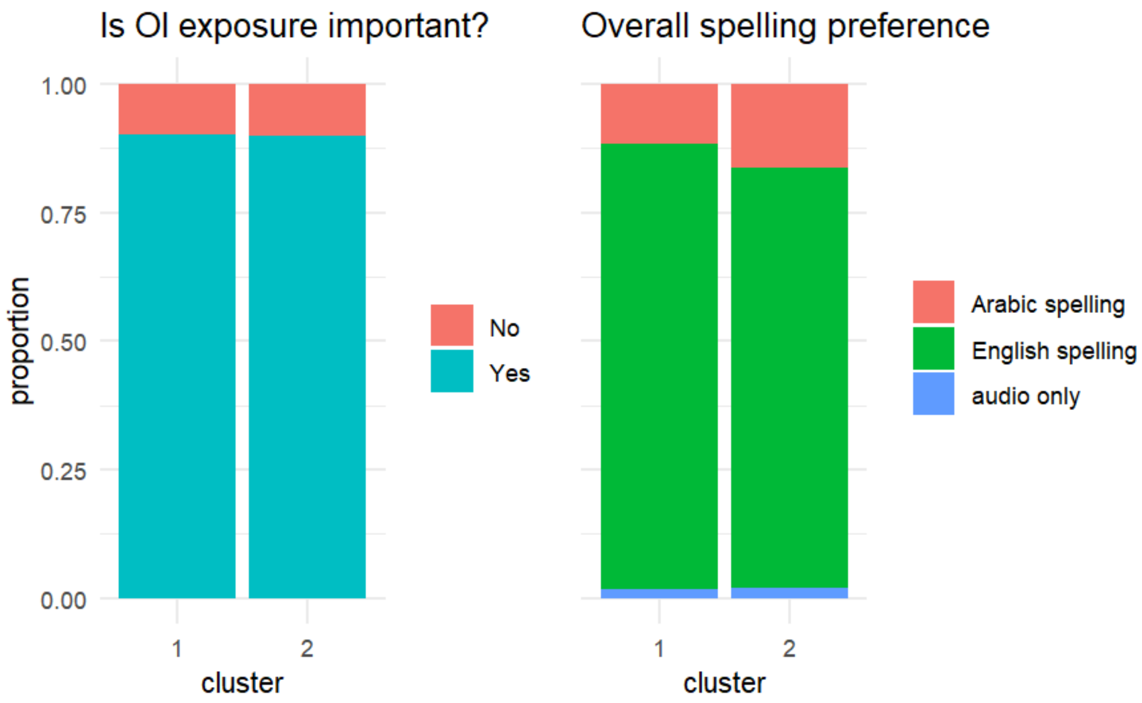


Figure 7.4: Bar plots of questionnaire responses related to OI importance and spelling preference, by proficiency cluster

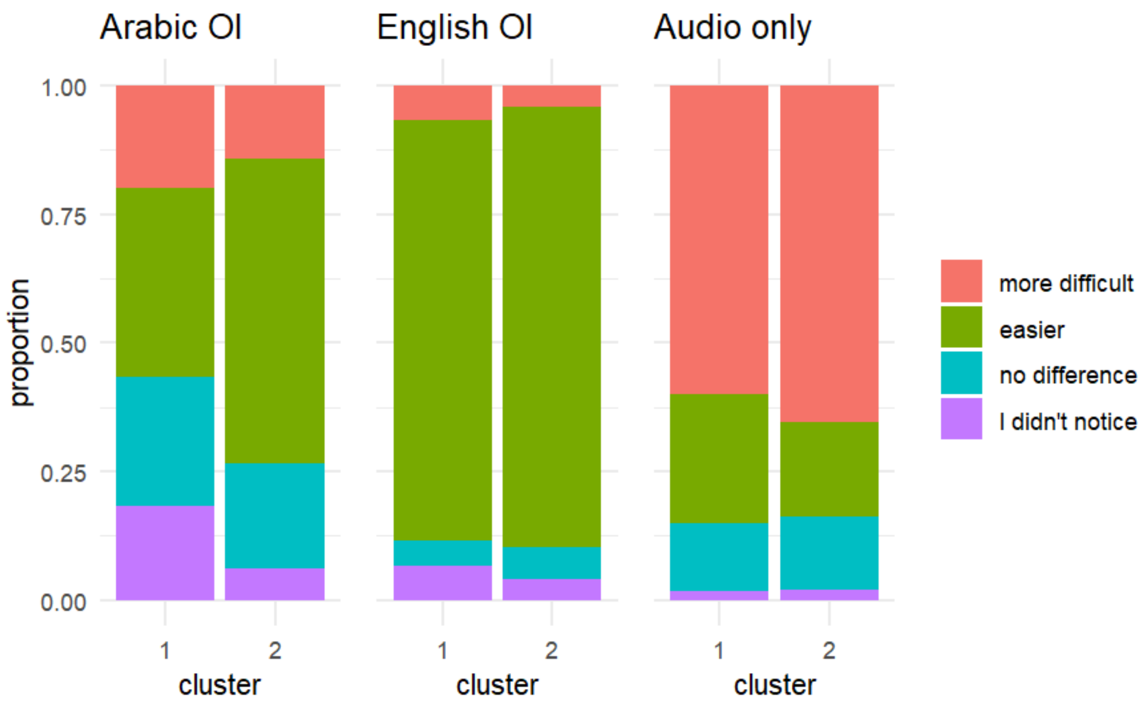


Figure 7.5: Bar plots of questionnaire responses related to OI condition difficulty, by proficiency cluster

7.3.3 Summary

The use of cluster analysis improves understanding of the dimensions of proficiency within the present study, demonstrating that L1 Arabic participants are best understood to fall within two higher and lower L2 English proficiency groups. Analysis through this lens confirms previous findings that higher proficiency learners perform more accurately in the matching task and lower proficiency learners are more likely to use strategies related to L1 Arabic literacy. It also reveals that the proficiency is less influential in relation to perceived influence of OI and beliefs about the importance of written input. For a closer examination of the nuances of proficiency in relation to the key variables of this study (i.e., performance, perceptions, and strategies), the next section discusses a sample of learner profiles from each of the clusters. Groupings are labelled as *lower* and *higher*, rather than *low* and *high*, as these clusters more likely reflect intermediate and advanced levels of proficiency. This is based on the background information provided by the participants and the exploration of the three-cluster solution, which indicates that within the lower proficiency cluster, there is a larger grouping of mid-level participants and a smaller subset of low-level participants.

7.4 Learner profiles by proficiency group

Individual differences have become a topic of great interest in SLA research, with good reason (Li et al., 2022). Adult language learning is defined by diversity; diversity of background, beliefs, starting point, journey, and destination, to name a few considerations. Thus far, this thesis has approached research questions by looking for broader patterns and points of unity within that diversity, through building predictive models, creating categories and reducing the dimensions of data. The small sample of learner profiles presented here intentionally embraces the variation and context that accompanies each participant in the study. This is with the aim of adding nuance to the findings discussed thus far, as well as highlighting points of interest for future research. As mentioned, these profiles are clustered into groups of lower and higher proficiency learners. Four profiles are presented from each proficiency cluster, where learning profiles (LP) 1-4 are from the lower proficiency cluster (*n*=49) and LP 5-8 are from the higher proficiency cluster (*n*=60). Profiles were purposively selected to highlight diverse backgrounds, performance in the study, and views provided in the posttest questionnaire.

Each profile presents demographic information, such as nationality, location at the time of completing the study, gender, age, level of completed education, and knowledge of additional languages. Additionally, English language experience is provided, including age of onset, self-reported level, proficiency test score, estimated proportion of daily exposure to English, time spent living in an anglophone country, and English language qualifications. Next, an overview of performance on the audio-visual matching task is given, with overall accuracy, mean *d′* scores, and mean response time. Also, mean *d′* scores are provided for /f-v/ contrast items in each orthography condition. Breaking down the scores for each individual in this way draws attention to the limited items and trials per condition and per person, yet still offers insight into orthographic effects. Environmental characteristics, such as device type, audio setup, and distractions, add further context.

The next section of the LP contains the speculative spellings provided for the audio-only target items. The four corresponding images were presented separately in a random order for each participant, who then typed a suggested spelling as part of the post-test questionnaire. Audio-only words were chosen to assess whether participants were generally able to perceive and encode the target contrasts, without written input. This was included as a rudimentary measure after the perception and production tasks were removed after piloting (chapter 3.2). Next, the open and multiple-choice responses from the post-test questionnaire are presented, including participant perceptions of OI, spelling preference, perceived influence of OI during the study, and perceived importance of OI in L2 word learning more broadly. The final section presents the LLS reported by each learner and their strategy component scores for *producing words*, *using Arabic OI for support*, *evaluating learning*, and *analysing sounds and letters*. The open responses in the last two sections also contain the tags applied during the qualitative content analysis. Thus, these profiles bring together the results from all three analysis chapters presented so far, with added context and individual differences included.

7.4.1 Lower proficiency learner profiles

The first three learner profiles present lower proficiency participants, who achieved the same number of correct answers in the matching task. While their performances are all broadly low, with accuracy just above chance, the *d′* scores for each OI condition reveal different influences of written input for each learner. The difficulty posed by the /f-v/ contrast items is further reflected in their speculative spelling, as none of these participants accurately produced the target items in writing. They all demonstrated some awareness of the need to distinguish between /f-v/ and /m-n/; however, this was not entirely successful. Aside from their performance on the task, these learners are united by their national context of Saudi Arabia. A large proportion of L1 Arabic participants in this study were from Saudi Arabia and performed with both high and low accuracy. Therefore, there is no suggestion of a relationship between national context and lexical encoding ability[48]. Additional shared characteristics include that they all began learning English between the ages of 10-12, declare their level to be either beginner or intermediate, speak no additional languages, and have completed at least an undergraduate degree.

Regarding the particulars of each learner, LP 1 (male, 22) has the lowest proficiency. He scores less than 50% on the proficiency test, English makes up less than 10% of his daily linguistic exposure, and he has not spent any time living in an Anglophone country. The *d′* scores for each OI condition of the matching task reveal lower accuracy encoding the /f-v/ contrast when words were accompanied with any written input. This aligns with the learner's perception that written input could distract from listening and the statement that it is not necessary to be exposed to written forms when learning new words. The second open response connects this view to the cognitive load associated with memorising the written and spoken form simultaneously. However, his preference was to see all the words written in English during the study. Interestingly, moderate accuracy in the no OI condition and awareness of the /f-v/ contrast in the speculative spelling imply that the learner can perceive the contrast, but its phonological representation may not be sufficiently robust to assess in tandem with written input. Looking at the reported strategies, there is little to note apart

[48] That being said, participants from North Africa, who speak Maghrebi varieties of Arabic, may have an advantage lexically encoding the /f-v/ contrast due to the acquisition and integration of French. However, the sample was inadequate for further exploration of this point.

from some reference to associative learning. Overall, this learner found the task difficult to navigate, particularly regarding the simultaneous exposure of written and spoken language.

Table 7.3: Learner profile 1: AR5450539

<table>
<tr><td>**Learner characteristics**
Nationality: Saudi
Location: Saudi Arabia
Gender: Male
Age: 22
Completed education: Bachelors
L3: none</td><td>**English Language experience**
Age of onset: 12
Level: beginner
Proficiency test score: 42%
English exposure: 8%
Time in L1 English country: none
English qualification: none</td></tr>
<tr><td>**Audio-visual matching task**
correct: 29 (58%), *mean d'*: 0.62, *mean RT*: 2299

/f-v/ with Arabic OI *d'* = 0.76
/f-v/ with English OI *d'* = 0.76
/f-v/ with no OI *d'* = 1.28</td><td>**Environmental characteristics**
Device: computer
Audio setup: device speakers
Distraction amount: 0%
Distraction type: device notification</td></tr>
<tr><td colspan="2">**Speculative spelling**
/famel/ - <fidos> /makem/ - <vidos>
/vamel/ - <macos> /nakem/ - <macom></td></tr>
<tr><td colspan="2">**Perception of OI**
OI condition preference: seeing English spelling

***Perceived OI influence*:**
في الغالب لا يحدث فرق بل ربما يشتت قليلا النظر الى الكلمة المكتوبة والصورة والاستماع
[*Often, it doesn't make a difference. It probably distracts a little looking at the written word and the image and listening* [DOESN'T HELP - DISTRACTING (FROM SOUNDS)]]

Perceived OI importance: no
لانه يصعب علي المهمة في حفظ الاملاء للاحرف وحفظ الكلمة المستمعة
[*Because it's a hard task for me to memorise the spelling of the letters and memorise the sound of the word* [IT'S TOO MUCH TO PROCESS]]</td></tr>
<tr><td colspan="2">**Language Learning Strategies**
Self-reported strategies:
ربط الاشكل وتشابه الكلمات
[*connecting the shape and similarity of words* [ASSOCIATING]]

Producing score: -0.01
Using Arabic OI score: 0.26
Evaluating score: 0.23
Analysing score: 0.40</td></tr>
</table>

In contrast LP 2 (female, 42) has higher proficiency, evidenced by a good score on the proficiency test, self-reported intermediate level, and an IELTS qualification of 5.5. Looking at her *d′* scores for each OI condition in the matching task, she has not been able to lexically encode the /f-v/ contrast, with notably low accuracy associated with Arabic OI. Despite this, she reports a preference for learning the words with the Arabic spelling, and suggests that both English and Arabic written forms help her to visualise words. She believes it is important to see the spelling of new words, mentioning the way she then makes personal connections with words she knows in Arabic. This is echoed in her reported strategy usage, where she gives a concrete example of an association between an object, presumably the blue /nakem/, with the Arabic word <نجم> [*najm*]. Evidence of this L1 transfer is then present in her speculative spelling of /nakem/ - <نجت> [*najt*], in addition to the fact that all the words are written in Arabic and lacking the distinction between /f-v/. She also scores relatively highly for strategies *using Arabic OI for support*, with a low score for *producing*. Altogether, this reveals a pattern of reliance on L1 Arabic language and literacy experience, and a lack of awareness of cross-linguistic interference.

Table 7.4: Learner profile 2: AR5357676

Learner characteristics *Nationality*: Saudi *Location*: Saudi Arabia *Gender*: Female *Age*: 42 *Completed education*: Masters *L3*: none	**English Language experience** *Age of onset:* 12 *Level:* intermediate *Proficiency test score:* 83% *English exposure*: 10% *Time in L1 English country:* none *English qualification*: IELTS 5.5
Audio-visual matching task *correct: 29* (58%), *mean d′*: 0.82, *mean RT*: 1999 /f-v/ with Arabic OI *d′* = -0.52 /f-v/ with English OI *d′* = 0 /f-v/ with no OI *d′* = 0	**Environmental characteristics** *Device:* computer *Audio setup:* device speakers *Distraction amount:* 0% *Distraction type:* none
Speculative spelling /famel/ - <فامل> [*faml*] /makem/ - <فدم> [*fadm*] /vamel/ - <مادت> [*madt*] /nakem/ - <نجت> [*najt*]	
Perception of OI	

***OI condition preference**:* seeing Arabic spelling

Perceived OI influence:

نعم فقد ساعدني في تذكر شكل الكلمة كثيرا سواء الكلمة بالعربية او الانجليزية فأنا أتخيلها واتذكر لو حتى حرفا واحدا

[*Yes, it helps me* [HELPS] *remember the shape of the word a lot, whether the word was in Arabic or English, I imagine it and I remember even if only one letter* [CONNECT AUDIO AND VISUAL INPUT - *VISUALISE WORDS*]]

Perceived OI importance: yes

يساعدني جدا في تذكرها وربطها بكلمة في لغتي الام ولو برابط مضحك خاص بي انا .. لا اخبره لأحد

[*it helps me a lot to remember it* [REMEMBERING] *and connect it to to words in my mother tongue if only with a funny link of my own* [VISUAL ASSOCIATION]*, I…don't tell it to anyone*]

Language Learning Strategies

***Self-reported strategie**s:*

احيانا من شكل الصورة احاول اربط شكل الصورة بشي عربي مثل الشي الازرق فقد ربطه بكلمة (نجم) بالعربية لان له أذرع

[*Sometimes from the shape of the image, I try to connect the shape of the image with something Arabic, like the blue thing I associated with the word (star) in Arabic because it had arms* [ASSOCIATING - *WITH SOMETHING FAMILIAR*]]

Producing score: -1.66
Using Arabic OI score: 1.09
Evaluating score: 0.26
Analysing score: -0.84

Turning to LP3 (female, 25), there is a notable difference in experience with English, as she scored highly on the proficiency test, reports higher daily English exposure, and has had the experience of living in an Anglophone country for more than a year. Regarding different OI conditions, English spelling appears to facilitate encoding of the /f-v/ contrast, in comparison to both Arabic OI and no OI. This aligns with her preference for seeing English spelling, on the basis that it supports the words in memory. Looking at her speculative spelling, she appears to overcorrect to the nonnative /v/ sound for both /f-v/ words, and struggles to remember the full word. She goes on to observe the inconsistency between spelling and pronunciation in English, but suggests that this is why early exposure to written forms are important. Presumably, this implies a high priority for L2 literacy and an awareness that spelling is not easily deduced from spoken forms, so must be explicitly taught. The main strategy she reported was *associating words with something familiar* and she had low scores for both *evaluating* and *analysing*. Based on analysis in chapter 4, it is likely that her low score relates to the amount of distraction she reports during the study. The mention of background noise may explain her score, in spite of her English language experience, as it is likely to interfere with her perception of the target contrasts in the study.

Table 7.5: Learner profile 3: AR5247103

<table>
<tr><td>Learner characteristics
Nationality: Saudi
Location: Saudi Arabia
Gender: Female
Age: 25
Completed education: Masters
L3: none</td><td>English Language experience
Age of onset: 10
Level: intermediate
Proficiency test score: 92%
English exposure: 32%
Time in L1 English country: more than 1 year
English qualification: Master's degree</td></tr>
<tr><td>Audio-visual matching task
correct: 29 (58%), mean d': 0.58, mean RT: 2331

/f-v/ with Arabic OI $d' = 0$
/f-v/ with English OI $d' = 1.28$
/f-v/ with no OI $d' = 0$</td><td>Environmental characteristics
Device: computer
Audio setup: earphones
Distraction amount: 62%
Distraction type: background noise</td></tr>
<tr><td colspan="2">Speculative spelling
/famel/ - <vadet> /makem/ - <macom>
/vamel/ - <va> /nakem/ - <naket></td></tr>
<tr><td colspan="2">Perception of OI
OI condition preference: seeing English spelling

Perceived OI influence: yes
It was easier to remember the words. [HELPS - REMEMBERING]

Perceived OI importance:
yes - In English , there is a difference between the speling and the pronouncing of words, which make it importing to learn the the speling of a new word when hearing it. [CLARIFY AND CONSOLIDATING NEW LANGUAGE - DECODING SOUNDS AND LETTERS]</td></tr>
<tr><td colspan="2">Language Learning Strategies
Self-reported strategies:
link it with somthing i know [ASSOCIATING - WITH SOMETHING FAMILIAR]

Producing score: 0.17
Using Arabic OI score: 0.89
Evaluating score: -1.16
Analysing score: -1.7</td></tr>
</table>

The final example from the lower proficiency cluster is LP 4 (female, 22) from Kuwait, who scores well on the proficiency test but self-reports an intermediate level with limited exposure to English in daily life. She has never lived in an Anglophone country and has no formal English qualifications. She performs very well on the matching task, with comparative accuracy across OI conditions. She also accurately spells the four target items that were presented without OI, demonstrating she has perceived and lexically encoded the /f-v/ contrast. Her spellings are given in both English and Arabic, which indicates an awareness of the cross-scriptal patterns of /f-v/ words. She reports no distractions during her participation and a preference for seeing the words accompanied by English spelling, although she does not think written input made a big difference to her learning. With regards to the broader importance of written input, she believes it facilitates the distinction between similar sounding words, implying this can be more reliable than listening. Her highest score for *producing* aligns with her open response, which details both clusters and strings of cognitive and metacognitive strategies. Despite her preference for English spelling and support for OI more generally, she mentions the distracting influence of written input. She also describes the ways she plans her learning and exposure to different elements of the input. Overall, this suggests a learner who is reflective, actively engaged in learning the target items, and has the ability to direct her attention to accurately encode the word forms.

Table 7.6: Learner profile 4: AR5730442

Learner characteristics *Nationality*: Kuwaiti *Location*: Kuwait *Gender*: Female *Age*: 22 *Completed education*: Bachelors *L3*: none	**English Language experience** *Age of onset:* 6 *Level:* intermediate *Proficiency test score:* 92% *English exposure*: 9% *Time in L1 English country:* none *English qualification*: none
Audio-visual matching task *correct: 44* (92%), *mean d':* 2.08, *mean RT*: 1364 /f-v/ with Arabic OI *d'* = 1.81 /f-v/ with English OI *d'* = 1.81 /f-v/ with no OI *d'* = 1.81	**Environmental characteristics** *Device:* mobile *Audio setup:* earphones *Distraction amount:* 0% *Distraction type*: none
Speculative spelling /famel/ - <Famo فامو> /makem/ - <Makom ماكوم>	

/vamel/ - <Vamo فامو> /nakem/ - <Nakom ناكوم>

Perception of OI
OI condition preference*:* seeing English spelling

Perceived OI influence**:**

لا يوجد فرق كبير

[*There wasn't a big difference* [DOESN'T HELP]]

Perceived OI importance**:** yes -

لأستطيع تمييزها عن الكلمات المتشابهة لأن الاستماع للغة الانكليزية غير واضح

[*To be able to distinguish it from the similar words because listening to the English language is not clear* [CLARIFYING AND CONSOLIDATING NEW LANGUAGE - *CORRECT COMPREHENSION*]

Language Learning Strategies
Self-reported strategies:

بالربط المنطقي بين الشكل والكلمة ، ولم أركز كثيرا على الحروف المتشابهة كي لا أتشتت كنت أتعامل مع كل كلمة على حدى
ركزت على الصوت، في بعض الأحيان كنت أتعمد تغطية الكتابة بأصابعي وأحاول نطق الكلمة قبل الصوت المسجل

[*With the logical connection between the shape and the word [*ASSOCIATING*], and I didn't focus much on the similar letters so I wouldn't be distracted. I dealt with each individual word focusing on the sound [*FOCUSING - *ON SOUNDS], sometimes I deliberately covered the writing with my fingers [*DIRECTING LEARNING*] and tried to pronounce the word before the recorded sound [*REPEATING*]*].

Producing score: 0.88
Using Arabic OI score: -0.71
Evaluating score: 0.74
Analysing score: 0.34

These four profiles, taken from a sample of 49 participants, offer a window into the individual differences within this proficiency cluster. Looking first at the proficiency levels of these participants, it again appears that this cluster better encompasses mid rather than low proficiency learners. Overall, these four profiles add context and nuance to factors already identified throughout the thesis, namely the potential for written input to distract from auditory input, as well as distinguish between similar sounds. The interfering influence of heavy reliance on L1 Arabic literacy was also reiterated, and is shown to persist into intermediate language learning. Finally, examples of broadly distracted and attentive participants have been given, where high linguistic awareness and task engagement is reflected in task performance and creative strategy usage. Therefore, while proficiency may broadly predict accurate lexical encoding of target contrasts, this demonstrates a more complex reality.

These profiles also provide insight into situations where participant reflections do and do not align with performance. For example, LP1 appeared aware of the added difficulty associated with attending to and encoding multimodal input, which was also reflected in lower *d′* scores compared to audio-only words. Even with this awareness, it appears he was not able to direct or suppress attention to mitigate this interference. LP4 also mentions a similar kind of distraction but appears able to direct her learning so that, as she reflects, there was little difference in her performance across all three OI conditions. Distraction was also a relevant factor for LP3, but this time in terms of her environment rather than the multimodal input. Her preference for English spelling aligned with her higher performance with these words compared to other OI conditions. Her higher English proficiency and exposure would suggest a more robust mental representation of /f-v/ is likely. Thus, the systematic representation of the contrast in English may have facilitated her discrimination, despite the noisy environment. Finally, LP2 seems unaware of the negative effects of relying on Arabic written input to encode the /f-v/ contrast. Despite her generally low scores, and particularly poor performance with Arabic OI, she reports that both English and Arabic spelling help her learn the words, with a preference of Arabic OI. The consistent reporting of active L1 transfer in relation to association and visualisation throughout her responses suggests a belief that this approach supports her learning, despite evidence to the contrary. The reasons behind this mismatch are discussed further in chapter 8. For now, individual differences are explored in relation to the higher proficiency cluster.

7.4.2 Higher proficiency learner profiles

As mentioned, the participants in the higher proficiency cluster have been shown to outperform the lower proficiency group, with regards to accuracy on the matching task. However, there were instances of low task performance in this group, as well. The four learner profiles presented for this cluster include two participants who struggled to perform above chance on the matching task, and two participants who performed well. All these learners began learning English between the ages of 7-11 and were between the ages of 26 and 32 when they participated in the study. The first three learners have completed a postgraduate degree and two report knowledge of at least one additional language. All state their English level to be advanced or near-native, with scores of 76-100% on the proficiency test, and all report high levels of exposure to English in their daily lives.

The first profile, LP5 (female, 26) from Saudi Arabia, responded quickly during the matching task, but with low accuracy. Looking at the *d′* scores in each OI condition, it appears she has not been able to successfully encode the /f-v/ contrast. Written productions of the words confirm this and her spelling more closely corresponds to the target items /fadet/ and /vadet/, which were presented with the Arabic OI spelling <فادت>. This may indicate interference effects from the L1 orthography, which she overcorrects by mapping both /f/ and /v/ to <v>. Her open responses further clarify that she was aware of her confusion, particularly in relation to Arabic OI. She reports a preference for learning the words accompanied by English spelling. Additionally, she believes that it is important to see the written forms when learning new words, as the association between sound and writing is beneficial, as is the opportunity to check the pronunciation of different sounds. This reiterates comments from LP4, that listening in the L2 is not always clear. Rather than focusing on distinguishing sounds, this learner implies that the letters offer a crutch for misperception. Overall, the information provided by this learner suggests that L1 OI interference can be a factor that persists into high levels of L2 proficiency and the benefits of the systematic English spelling may not always extend to confusable L2 contrasts. The low accuracy encoding the /f-v/ contrast across all OI conditions, including audio-only and without any reported distractions, indicates that this contrast was not sufficiently perceptible to the learner. Alternatively, reported confusion between minimal pair items and low use evaluating strategies may be relevant considerations.

Table 7.7: Learner profile 5: AR5336601

Learner characteristics	**English Language experience**
Nationality: Saudi *Location*: Saudi Arabia *Gender*: Female *Age*: 26 *Completed education*: Masters *L3*: Turkish	*Age of onset:* 11 *Level:* advanced *Proficiency test score:* 100% *English exposure*: 28% *Time in L1 English country:* none *English qualification*: IELTS 6.5

<table>
<tr><td>Audio-visual matching task
correct: 27 (56%), mean d': 0.39, mean RT: 1356

/f-v/ with Arabic OI d' = -0.76
/f-v/ with English OI d' = 0
/f-v/ with no OI d' = 0</td><td>Environmental characteristics
Device: mobile
Audio setup: headphones
Distraction amount: 1%
Distraction type: none</td></tr>
<tr><td colspan="2">Speculative spelling
/famel/ - <vadit > /makem/ - <nakim>
/vamel/ - <vadit > /nakem/ - <makim></td></tr>
<tr><td colspan="2">Perception of OI
OI condition preference: seeing English spelling

Perceived OI influence:
writing words in Arabic made it somehow challenging to learn them. [DOESN'T HELP]

Perceived OI importance: yes -
to draw a connection between the sound and the written form [VISUAL ASSOCIATION]. it also helps if I miss how each sound is pronounced. [CLARIFYING AND CONSOLIDATING NEW LANGUAGE - CORRECT PRODUCTION]</td></tr>
<tr><td colspan="2">Language Learning Strategies
Self-reported strategies:
i got confused if I should write N or M in both makim and nakim and likewise with vadit and fadit. /v/ & /f/ [EXCL - NO STRATEGY]

Producing score: 0.80
Using Arabic OI score: -0.13
Evaluating score: -1.22
Analysing score: 0.74</td></tr>
</table>

Also from Saudi Arabia, LP6 (female, 28) achieved the same score as the first three profiles in the lower proficiency group, with a relatively fast response time. She reported moderate distraction during the experiment, due to interruption from a member of her household, but her speculative spelling reveals that she was able to lexically encode both the /f-v/ and /m-n/ contrasts for the words presented without any OI. However, *d'* scores in each orthography condition revealed an inhibitory influence of OI, particularly Arabic spelling. She reports knowledge of three additional languages, which indicate that, as well as being multilingual, she may be multiliterate in four different writing systems (Arabic alphabet, Roman alphabet, Hangul alphabet and Japanese kanji, hiragana, and katakana). This is speculative as her oral and written proficiency in these languages is unknown. However, her knowledge of French may be relevant for encoding the /f-v/ contrast items. She reports a preference for seeing the English spelling when learning the words, although she also mentions the potential for written forms to distract from other input, specifically the picture. Further to this, she believes that it is not important to see written forms when learning new words, referencing the lack of systematic grapheme-phoneme correspondences in certain writing systems. This is likely related to previous language learning experience, as French and English can be highly inconsistent in their GPC rules, and logographic Japanese kanji GPCs are opaque, in contrast to hiragana and katakana syllabaries. With regards to LLS, her scores were generally low. She reports repetition and various forms of association. However, she was not convinced that repeating supported her learning and appears to have mostly focused on the images, perhaps at the expense of the word forms.

Table 7.8: Learner profile 6: AR5336140

Learner characteristics *Nationality*: Saudi *Location*: Saudi Arabia *Gender*: Female *Age*: 28 *Completed education*: Masters *L3*: Korean, Japanese, French	**English Language experience** *Age of onset:* 10 *Level:* advanced *Proficiency test score:* 75% *English exposure*: 33% *Time in L1 English country:* none *English qualification*: Master's degree Applied Linguistics
Audio-visual matching task *correct: 29* (58%), *mean d':* 0.70, *mean RT*: 1823 /f-v/ with Arabic OI *d'* = -0.76 /f-v/ with English OI *d'* = 0 /f-v/ with no OI *d'* = 1.05	**Environmental characteristics** *Device:* computer *Audio setup:* earphones *Distraction amount: 28%* *Distraction type:* member of household
Speculative spelling /famel/ - <famel? > /makem/ - <makum?> /vamel/ - <vadel...probably> /nakem/ - <nakum>	
Perception of OI ***OI condition preference**:* seeing English spelling ***Perceived OI influence***: sometime i get distracted by the written word as it does not make me focus as much on the picture. [DOESN'T HELP - DISTRACTING (FROM SOUNDS)]	

Perceived OI importance: no - beacuse sometimes writing does not refelct how we speak the language. there may be little correspondnce between the writing systme and the sound of a language. [SPELLING DOESN'T REFLECT PRONUNCIATION]
Language Learning Strategies ***Self-reported strategies****:* sometimes repeating the word [REPEATING] but not always(this failed me) [EVALUATING] + creating an image for the pic+word... ex the word famis for me seemed like a famous person who held his\her hand high.. trying possibly to wave to the fans? [ASSOCIATING - WITH SOMETHING FAMILIAR]] other than that, some images were easier to remember because of the color(i.e., the orange thingy). [ASSOCIATING - WITH VISUAL DETAIL] *Producing score:* -0.07 *Using Arabic OI score:* -0.46 *Evaluating score:* -0.87 *Analysing score:* -0.40

The next profile is from another Saudi Arabian learner, LP7 (female, 32), who scores highly on the matching task. Also, her speculative spelling reflects accurate lexical encoding of the target phonological contrasts without OI. However, a closer look at her accuracy in different OI conditions reveals that encoding of this contrast was inhibited by exposure to Arabic OI but promoted by English OI. This learner has spent more than a year living in an Anglophone country and almost half of her daily linguistic exposure is reportedly in English. She states a preference for learning new words accompanied by the English spelling but notes that she focused more on the sound and image during the task. More generally, she believes that OI is important when learning new words, as a guide for pronunciation. Her strategy usage then reveals a low score for *producing* and high score for *analysing* during the task. This is expanded upon in her open response, where she details associating strategies based on her linguistic knowledge in both Arabic and English. This quote was previously discussed as an example of L1 interference in chapter 6, where the learner perceives the /f-v/ contrast but intentionally maps both lexical items to a single word in Arabic. When examined in the context of the broader learner profile, she appears able to harness her knowledge of both languages, and the relationship between them, in a way that supports the accurate encoding of the target forms without written input and with English

OI, but not Arabic OI. This may be an instance where her representation of the contrast is sufficiently robust to take advantage of systematic-distinct English OI, yet interference persists in the context of incongruent-shared Arabic OI.

Table 7.9: Learner profile 7: AR5336722

<table>
<tr><td>Learner characteristics
Nationality: Saudi
Location: Saudi Arabia
Gender: Female
Age: 32
Completed education: Masters
L3: none</td><td>English Language experience
Age of onset: 7
Level: near-native
Proficiency test score: 92%
English exposure: 47%
Time in L1 English country: more than 1 year
English qualification: Master’s degree</td></tr>
<tr><td>Audio-visual matching task
correct: 42 (88%), mean d’: 2.07, mean RT: 2251

/f-v/ with Arabic OI   d’ = 0
/f-v/ with English OI   d’ = 2.56
/f-v/ with no OI   d’ = 1.81</td><td>Environmental characteristics
Device: computer
Audio setup: earphones
Distraction amount: 0%
Distraction type: none</td></tr>
<tr><td colspan="2">Speculative spelling
/famel/ - <Famo>   /makem/ - <Makom>
/vamel/ - <Vamo>   /nakem/ - <Nakom></td></tr>
<tr><td colspan="2">Perception of OI
OI condition preference: seeing English spelling

Perceived OI influence:
Maybe it has made it easier a little bit, but I relied more on memorizing the shape and the word heard [DOESN’T HELP - BETTER TO IGNORE/ FOCUS ON SOUNDS].

Perceived OI importance: yes-
To realize how it's pronounced later when I see the word. [CORRECT PRODUCTION]</td></tr>
<tr><td colspan="2">Language Learning Strategies
Self-reported strategies:
Connecting the shape to something in real world. e.g. those that are called 'famo' or 'vamo', I try to connect the word heard which is a bit similar to "فم" in Arabic saying the shape of the those pictures can be entered inside the 'mouth=فم". Another example, the shape that's called 'famis' ...I say it looks like a unique flower..so I remind myself that it's famis "like or driven from the English word famous". So in short, try to connect the new words to known words in either language Arabic or English. [ASSOCIATING WITH SOMETHING FAMILIAR]</td></tr>
</table>

Producing score: -1.55 *Using Arabic OI score:* 0.18 *Evaluating score:* 0.18 *Analysing score:* 1.18

Finally, LP 8 (male, 29) is a Palestinian learner, who was located in Jordan at the time of the participation. He has never spent any time in an Anglophone country but achieves a high proficiency test score, and estimates around 50% of his daily linguistic exposure is in English. He gained a moderately high score in the matching task; however, he reported some distraction during participation, due to background noise. The issue of background noise is potentially exacerbated by using mobile phone speakers, rather than headphones. Based on his speculative spelling, there is evidence that he was able to perceive and lexically encode the /f-v/ contrast. However, *d'* scores for the different OI conditions reveal that he was only consistently accurate when the words were presented with English spelling. An awareness of this may be reflected in the preference for learning words accompanied by English spelling. He also mentions written forms helping to distinguish between similar sounding words and associative support when storing the words in memory. The focus on associations was echoed in the beliefs about written input importance, as well as LLS usage. He then scored highly on *producing* and *evaluating* strategies, while scoring low on *using Arabic OI for support*. Overall, this learner made use of English written input to lexically encode the confusable contrasts, with moderate success. The distraction of background noise may relate to a similar effect found with LP3, where accuracy was highest with English spelling and low for both other conditions. Thus, in the context of noise, the systematic spelling may enhance the salience of a contrast which has a robust mental representation. The influence of OI in noisy L2 spoken language processing could be another interesting avenue for further research, as well as the relevance of increased lexical competition across languages in noisy conditions (Guediche et al., 2023).

Table 7.10: Learner profile 8: AR5229722

Learner characteristics	**English Language experience**
Nationality: Palestinian *Location*: Jordan *Gender*: Male	*Age of onset:* 10 *Level:* advanced *Proficiency test score:* 100%

<table>
<tr><td>Age: 29
Completed education: Bachelors
L3: none</td><td>English exposure: 49%
Time in L1 English country: none
English qualification: none</td></tr>
<tr><td>Audio-visual matching task
correct: 34 (77%), mean d': 1.32, mean RT: 2892

/f-v/ with Arabic OI   d' = 0
/f-v/ with English OI   d' = 1.81
/f-v/ with no OI   d' = -0.75</td><td>Environmental characteristics
Device: mobile
Audio setup: device speakers
Distraction amount: 15%
Distraction type: background noise</td></tr>
<tr><td colspan="2">Speculative spelling
/famel/ - <Famot>   /makem/ - <makom>
/vamel/ - <vadot   /nakem/ - <makit></td></tr>
<tr><td colspan="2">Perception of OI
OI condition preference: seeing English spelling

Perceived OI influence:
As some of the words sound similar, seeing it wtitten helped me [HELPS] distinguish between them when hearing them so I would say it helped a lot in that case [CLARIFY WHAT I HEARD - DIFFERENTIATE BETWEEN SIMILAR SOUNDS/WORDS], it also helped with other cases as linking the image and the sound with a written word made it easier to remember [CONNECT AUDIO AND VISUAL INPUT - ASSOCIATE WORD AND IMAGE].

Perceived OI importance: yes -
This way I got more information which means more links to the word to easily recall it. [VISUAL ASSOCIATION]</td></tr>
<tr><td colspan="2">Language Learning Strategies
Self-reported strategies:
I tried to create a linkage between how the word sounds and how the image looks. [ASSOCIATING]

Producing score: 1.54
Using Arabic OI score: -1.43
Evaluating score: 1.63
Analysing score: 0.86</td></tr>
</table>

These four profiles, taken from a sample of 60 participants in the higher proficiency cluster, exemplify additional context relevant to the successful lexical encoding of L2 phonological contrasts. As with the lower proficiency group, there are examples of varied performance, with even high proficiency participants struggling to accurately encode the difficult /f-v/ contrast, as was noted in the analysis presented in chapter 4. These profiles demonstrate individual differences in what is perceived to be distracting or supportive, as well as apparent contradictions between learner performance, preferences, and beliefs about the importance of written forms. For example, LP6 believes that OI is not necessarily important when learning new words, and was at times distracting during the task, yet articulates a preference for seeing English spelling rather than no written input. Both LP5 and LP6 refer to distractions and unreliability of written forms, where LP5 points more to L1 interference. LP7 also reports cross-linguistic transfer but appears to be able to make use of her linguistic knowledge in a way that supports, rather than undermines, accurate encoding of target contrasts with both English OI and no OI. Potentially, LP7 is more proficient, with more robust phonological representations than LP5, which invites the question of whether there is a point in a learners' proficiency trajectory where cross-linguistic transfer becomes something that can be harnessed without undue interference. That being said, even LP7 experiences difficulty encoding /f-v/ contrast words when accompanied by Arabic OI. These profiles also add useful context to how learners understand the role of written input in associative learning and demonstrate the connection between written input and language production for learners, where spelling is perceived to be a guide for pronunciation.

7.4.3 Summary

The eight learner profiles presented here have demonstrated that individual differences and context can reveal useful insights into proficiency and other factors relevant to task performance and participant perceptions. Both lower (LP1-4) and higher (LP5-8) proficiency clusters included learners who struggled to encode the target /f-v/ contrast, and others that performed well in the matching task. This demonstrates that proficiency does not fully account for the patterns and variation in responses reported so far. Those with lower performance (1) provided little evidence they were able to perceive or lexically encode the confusable L2 contrast (LP2, LP5), (2) reported explicitly relying on L1 written input, which did not visually distinguish the L2 target contrast (LP3), or (3) indicated distraction during participation, either from the multimodal input (LP1, LP6) or background noise (LP2, LP8). Meanwhile, higher performing participants provided evidence of lexically encoding the difficult contrast and reported minimal influence of written input (LP4, LP7). Additionally, by taking a closer look at *d′* scores for each OI condition, there is more opportunity to untangle some of the discord between participant performance and perceptions. For example, LP1, LP4 and LP6 all mention that written forms can introduce distraction, in relation to cognitive load or unreliable sound-spelling relationships. However, each of the learners also reported a preference for learning with English spelling and appeared to manage distractions, successfully or unsuccessfully, in different ways.

The learner profiles presented also draw attention to the difficulty around classifying the proficiency of learners in this study. For example, the inclusion of both LP1 and LP3 in the same lower proficiency cluster seems problematic. The English experience of LP1 aligns with that of a beginner learner, while LP3 has a high proficiency test score, completed a master's degree in English, spent over a year in an Anglophone country, and around a third of daily linguistic exposure is in English. This highlights the fact that the higher proficiency group included participants who had even more English experience than LP3, as well as the wider range of lower English proficiency learners, noted when clustering the learners in section 7.3. Despite proficiency differences, both learners achieved low accuracy scores in the matching task. The value of looking at individual differences and context to participation is that the low performance of LP3 appears likely related to background noise, rather than proficiency. The relevance of environmental factors, such as background noise, arises again with LP8 in the higher proficiency group. It is of critical importance that participants are in a quiet location when participating in a study which involves learning words that differ by perceptually difficult sounds. Thus, clear gains to be made from including distraction checks in similar internet-based studies, and perhaps inclusion criteria should have been stricter.

Across both groups, interference from L1 orthographic input persists across all levels of proficiency. Even those with the highest level of English experience appear better able to take advantage of the systematic English OI, but unable to suppress interference from incongruent Arabic OI (LP7). Awareness of this inhibitory effect varied, where LP5 noted that Arabic spelling was more challenging, whereas LP2 reportedly embraced heavy reliance on L1 Arabic language and literacy knowledge. Despite their proficiency differences, both in fact performed very similarly on the matching task. Even though LP5 had an awareness of Arabic OI introducing an additional challenge, which was not evident with LP2, this appears not to have been insufficient to mitigate interference from shared-incongruent spelling.

Overall, these profiles highlight that there is much diversity in learner backgrounds, beliefs, and approaches to present study. Additional contextual information helps to understand the influence of proficiency and contributing factors to participant performance. Taking a more in-depth look at proficiency, understanding is improved of different forms of distraction during participation, both in terms of environmental factors and related to the experimental items themselves. It is also clear that, while participants are aware of the influence of orthography in the present study, they respond to that influence in different ways. The evidence presented here offers motivation to look more closely at cognitive control, cross-linguistic activation and competition, and the ability to perceive the target contrasts, in relation to proficiency and orthographic influence. The information in the learner profiles also adds to evidence that explicit LLS vary across learners, but bear little relation to word learning performance with the confusable contrasts.

7.5 Concluding remarks

The clustering of participants and examination of learner profiles in relation to proficiency has added clarity and depth to findings reported thus far, as well as raising questions for further research. First, it appears that proficiency in the present study is best understood in terms of two higher and lower proficiency clusters, which mostly correspond to advanced and intermediate learners, respectively. The relationships between proficiency, matching task performance, and reliance on L1 Arabic literacy, reported in previous chapters, was confirmed when looking at the participant clusters, including more measures of proficiency. The learner profiles then provided insight into individual differences, demonstrating high and low performance in both proficiency clusters, in relation to experimental, cognitive, and environmental factors. The more in-depth view into individual learner perspectives also highlighted varying awareness of cross-scriptal orthographic influence and ability to mitigate interference effects.

In line with previous studies, both intermediate and advanced learners articulated a clear preference for exposure to English script input, rather than the Arabic script, which may reflect what learners have become accustomed to with increasing proficiency (Hao & Yang, 2021). However, this deserves further attention in terms of the developmental point at which learners begin to view distinct script input as something familiar, or more congruent with broader linguistic exposure and supportive of learning. There is also evidence that participants who performed poorly in the task, irrespective of proficiency, may have been more biased by written input, in the sense that they were more affected by its presence (Mok et al., 2018). This orthographic bias could include a distracting influence, either related to cognitive load or cross-linguistic transfer. Indeed, this may connect to predictions made by Veivo and colleagues (2018; 2013) that the facilitative effects of OI are more evident in advanced learners, due to more robust phonological representations and heightened cognitive control of cross-linguistic information. However, there is also evidence that when there is incongruence in shared script input, interference effects persist, regardless of proficiency (Escudero et al., 2014).

Future research would benefit from including more ID measures relating to phonological awareness, auditory processing and working memory, as these are highly relevant to understanding participant perceptual and word learning abilities. It would also be enlightening to explore proficiency in relation to participants with different literacy profiles, including multi-script literacy and limited L1 literacy. Additionally, the persistence of orthographically-induced interference effects into high levels of proficiency, and limited evidence of facilitative effects with any OI when confusable contrasts are not well represented phonologically, calls into question early reliance on written input with lower proficiency learners. However, as has been shown, vocabulary and pronunciation learning involving less written input would likely prove unpopular with learners, for a variety of reasons. The next chapter draws together the findings presented in each of the analysis chapters, with further discussion of the theoretical implications, pedagogical applications, and proposals for future research.

Chapter 8: Discussion

8.1 Introduction

The primary research question laid out at the start of this thesis was:

Does exposure to the Arabic and English script differentially influence L1 Arabic speakers' acquisition of new words that differ by a difficult nonnative phonological contrast?

The short answer to this question, as evidenced by the analyses presented in chapters 4 to 6, is both yes and no. To understand this contradiction, it is helpful to return to the diagram presented in the introduction (figure 8.1), highlighting the relationship between task performance, participant perceptions and language learning strategies (LLS). For example, analysis of accuracy in the audio-visual matching task indicated that there was not a significant difference in performance depending on whether /f-v/ words were taught with Arabic or English spelling. Both had an inhibitory effect compared to no orthographic input (OI). Additionally, there was no evidence that script-specific or orthographically-influenced LLS had a significant influence on performance either. Thus, quantitatively, but perhaps superficially, there is little evidence that Arabic and English orthographic exposure had distinct influences on lexically encoding confusable L2 contrasts.

On the other hand, qualitative analysis of open responses to the post-test questionnaire clearly demonstrated that participants were aware of script-specific influences during word learning, in and beyond the experimental task. Participants articulated an awareness of varied orthographic influence in relation to lexically encoding the target contrasts in memory, beliefs about language learning more broadly, and the use of different LLS. The discord between participant perceptions about their learning and their ability to accurately encode a difficult L2 contrast reflects the complexity of the processes involved, but also the importance of understanding the motivating factors upon which participants consciously and unconsciously direct their learning. Input is not perceived or processed the same way by each learner, but rather influenced by literacy in known languages, cross-linguistic transfer, and learning goals. Additionally, looking more closely at the quantitative findings, proficiency level is crucial to understanding the varied influence of Arabic and English written input, both with regards to performance and strategy usage.

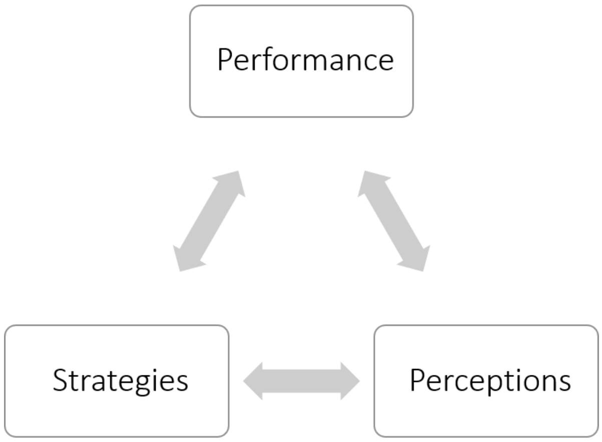


Figure 8.1 Relationship between core learner variables in the present study

The following sections of this chapter outline the key findings in more detail and in specific reference to the secondary research questions. These findings are then explored in relation to broader points of discussion emerging from the thesis. Firstly, theoretical contributions are presented, including a preliminary account of the minimal facilitation of orthographic information on L2 phonological acquisition, with specific reference to cross-scriptal written input. The relevance of error-driven learning and extensions of speech learning models are also discussed. This is followed by pedagogical recommendations, based on the outlined key findings and theoretical perspectives, where a modified version of the *Phonology and Orthography Language Learning Strategies* (POLLS) inventory is presented, alongside classroom-based applications. Next, methodological considerations are discussed, outlining strengths, limitations, and recommendations for the field. The chapter concludes with a summary of future directions.

8.2 Key findings

The analysis presented in each of the chapters from 4-6 sought to address one of the secondary research questions, stated initially in chapter 2.4. Table 8.1 provides a summary of the research questions and associated key findings, followed by further detail in subsections 8.2.1-3. Additionally, the findings from chapter 7 are summarised, which are not presented in relation to a separate research question, but take a closer look at the recurring influence of proficiency reported throughout the thesis. Thus, an adaptation of the primary research question, including the influence of proficiency, is included in the summary table 8.1.

Table 8.1 Summary of research questions and key findings

Chapter	Research question(s)	Key findings
4	To what extent do different types of orthographic input (OI) influence the lexical encoding of L2 phonological contrasts in memory during novel word learning?	➢ Accuracy encoding /f-v/ contrast items was lower with both Arabic and English OI compared to audio-only presentation. ➢ Accuracy improved with proficiency, yet even high proficiency participants struggled to perform above chance with /f-v/ items.
5	What influence of OI were learners aware of during word learning? a) What was the perceived influence of OI during the present study? b) What were the participants' beliefs about the importance of OI when learning new words more broadly?	Most participants: ➢ Were aware of orthographic influence, where effects differed depending on the script input. ➢ Expressed a clear preference for learning with English spelling. ➢ Perceived OI to be supportive of learning during the study. ➢ Reported that early exposure to written input was important when learning words in a new language.
6	What is the relationship between learners' language learning strategies and lexical encoding of L2 phonological contrasts in novel words?	➢ Participants drew on a range of cognitive and metacognitive strategies, involving phonological and orthographic knowledge. ➢ The only strategy which (marginally) predicted improved performance was *evaluating learning.* ➢ Lower proficiency participants reported higher reliance on *using Arabic OI for support.*
7	To what extent does proficiency mediate the effect of exposure to Arabic and English OI on L1 Arabic speakers' acquisition of new words that differ by a difficult L2 phonological contrast?	➢ Participants cluster into two groups of higher and lower proficiency learners. ➢ Confirmation that higher proficiency learners outperform lower proficiency learners and rely less on L1 OI. ➢ More attention needed on individual differences beyond proficiency.

8.2.1 Influence of written input on lexical encoding of contrasts

Chapter 4 responded to the first question by conducting analysis with mixed-effects models of accuracy and response time data from the word learning phase, followed by accuracy data from the audio-visual matching task. Analysis was conducted separately for the L1 Arabic learners of L2 English (n = 114) and a control group of L1 English-speakers (n = 118). Both groups completed the word learning task with mean scores over 90%, demonstrating that all could recognise the 12 target items, in the context of non-minimal pair distractors, with relative ease after a short learning phase. The key point of interest was whether participants had encoded the novel items with sufficient detail to accurately distinguish between minimal pair items, differing by /f-v/ (difficult) and /m-n/ (easy) phonological contrasts, in the matching task. L1 English data was analysed first, to assess the method and ensure that it was possible to encode the target contrasts with above chance accuracy. This was evidenced in the results, although there was unexpected difficulty with the /f-v/ contrast words in comparison to the /m-n/ words, suggesting that the perceptual salience of the contrast was generally lower. L1 English participants were more accurate when they had been presented with English spelling compared to Arabic spelling. However, focusing on the less salient /f-v/ contrast words, there was an interaction with OI, where accuracy was lower with words taught with any OI compared to audio-only. Overall, the above chance accuracy indicated that participants were able to lexically encode the target contrasts, although low perceptual salience reduced the benefit of L1 shared script input.

The subsequent analysis of L1 Arabic accuracy revealed main effects of phonological contrast and proficiency, where accuracy was significantly lower with /f-v/ words but improved as L2 English proficiency increased. Like the L1 English group, accuracy was higher when words were presented with English spelling compared to Arabic spelling. However, a significant interaction between OI and phonological contrast again revealed that accuracy was lower for /f-v/ words that were presented with any form of OI, compared to audio-only. In other words, any OI inhibited accurate encoding of items differing by the confusable /f-v/, whereas accuracy was supported by any OI when items differed by the easier /m-n/ contrast. Furthermore, while accuracy improved with increased proficiency, and in contrast to the L1 English group, even the highest proficiency participants struggled to perform above chance with words differing by /f-v/ accompanied by any written input.

These results suggest orthographic influence is mediated by the perceptual difficulty of the target contrast. Even when OI is familiar and letters map transparently to sounds, it is not always useful in visually disambiguating the speech signal, but rather enhances only what is already perceptible to the learner.

8.2.2 Participant perspectives on written input

To further unpack these findings, chapter 5 explored what orthographic influences participants were aware of during the experimental tasks, as well exploration of broader language learning beliefs in relation to written input. Open and closed responses to the post-test questionnaire were analysed using descriptive frequencies and qualitative content analysis. Most participants in both language groups reported being aware of exposure to written input and expressed a clear preference for learning the words accompanied by English spelling. Overall, participants in both language groups perceived written input to have an overwhelmingly positive influence on word learning, in and beyond the present study. Most participants suggested that written input supported the clarification and consolidation of new language in memory, including distinguishing difficult to perceive sounds and visual analysis of the words. These findings demonstrate that participants' perspectives are at odds with their performance in the matching task, specifically with the difficult /f-v/ contrast items. However, in both language groups, some participants also reported that it was better to focus on sounds, as OI could be distracting and introduce issues around cognitive load. These views were notably more present in the L1 English responses, who additionally mentioned the unfamiliarity of the Arabic script.

Regarding the distinction between the two scripts, most participants responded broadly without specifying the script. Some participants, in both groups, mentioned script specific differences, including awareness of the systematicity of grapheme-phoneme correspondences in both English and Arabic. In some cases, English spelling was seen to be more reliable than the Arabic spelling, regarding experimental items, while others noted English is generally more inconsistent. Overall, when participants compared the influences of the different scripts, there were varied perceptions around what was beneficial or not. In order to contextualise these responses, analysis of broader beliefs revealed that participants had differing conceptualisations of the relationship between sounds and symbols, ideas of "correctness" and learning goals. The importance of exposure to written input was particularly strong in the L1 Arabic group, where participants emphasised the use of OI to support correct production and comprehension of new vocabulary. These ideas were also present in the L1 English responses; however, a larger proportion of participants were critical of the importance of OI, based on previous language learning experience, concerns about cognitive load, and the belief that spelling does not reflect pronunciation. These results draw attention to the rich and varied ways in which language learners bring their knowledge of, and ideas around, language(s) to navigate the processing of multimodal input when learning new words. While there are unifying patterns across both language groups, the groups differ in the extent to which they perceive written input to be reliable, as well as the L1 Arabic group emphasising the benefits for memory and "correctness" of the form.

8.2.3 Participant strategies and written input

The third research question, addressed in chapter 6, sought to better understand both beliefs and performance by exploring the LLS that participants consciously employed during the experiment. This question was addressed with the post-test questionnaire, where open responses were examined through the lens of qualitative content analysis, to find out what LLS the participants remembered using during the experimental task. Quantitative analysis of Likert-scale responses to the POLLS inventory then offered complementary insight into: the frequency of phonological and orthographic LLS; underlying dimensions to the data; and the relationship between strategy usage, performance, and proficiency. Both qualitative and quantitative analysis demonstrated that participants used a range of cognitive and metacognitive LLS, which included strategies drawing on orthographic and phonological knowledge. For example, participants across both groups reported strategies of focusing on word-initial sounds, connecting sounds and letters, visualising English spelling, and repeating words. Affective influences were also noted, including feeling more confident when the English spelling had been presented. Examples highlighted the preference for strategies involving English written forms, particularly for the L1 Arabic group. Open responses further revealed that participants went beyond the input they had been exposed to and drew on broader literacy experience to form mental speculations of spellings.

Responses to the POLLS inventory were further analysed using principal component analysis, resulting in the extraction of four key strategy components for the L1 English and L1 Arabic groups. Within both groups, their components mapped to *analysing sounds and letters,* and *using Arabic written input for support.* Notably, the latter in the L1 English group reflected low use of strategies relating to Arabic written forms. The L1 English components also included *connecting words to meaning* and *using English written input for support*. Meanwhile, the L1 Arabic components included *producing new words* and *evaluating learning.* The L1 Arabic group did make use of strategies related to English written forms, but they were spread across different components, including *analysing sounds and letters* and *producing new words*. These findings offer further evidence that participants make use of both phonological and orthographic strategies, which are often interwoven, as well as demonstrating script-specific differences. Regression analyses investigating the relationship between strategy component scores, accuracy in the matching task, and L2 English proficiency of the L1 Arabic group revealed two further insights. Firstly, strategy usage did not predict improved accuracy, although there was marginal evidence that *evaluating learning* predicted higher accuracy. Secondly, proficiency predicted *using Arabic written input for support*, where higher English proficiency predicted lower reliance on Arabic OI.

8.2.4 Beyond proficiency

Chapter 7 investigated the extent to which L2 English proficiency related to the influence of Arabic and English script input on the acquisition of new words differing by a confusable L2 phonological contrast. Analyses in previous chapters relied on the score from a short test within the study to measure L2 English proficiency, whereas the cluster analysis presented in this chapter included multiple measures to group L1 Arabic participants by proficiency. This analysis led to the formation of a higher (advanced-near-native) and lower (beginner-intermediate) proficiency group. Descriptive and inferential analysis with the data clustered in this way confirmed previously reported patterns, where higher proficiency learners were more accurate in the matching task, and lower proficiency learners relied more on Arabic written forms as a strategy to support word learning. The results from chapter 5 had not previously been discussed in relation to proficiency, thus the clusters also proved useful in illustrating that spelling preference and beliefs about the importance of early exposure to written input did not differ by proficiency, apart from confirming a more frequent preference for Arabic OI in the lower proficiency group.

Subsequently, four learner profiles were presented from each proficiency cluster, including demographic characteristics, English language experience, performance in the matching task, and an overview of post-test questionnaire responses. These profiles highlighted individual differences between participants who achieved similar accuracy score but were affected by exposure to written input in different ways. Additionally, high and low-scoring participants were presented from both proficiency clusters, as well as participants with varying perceptions of orthographic influence on their learning. Attention was drawn to the importance of perceptual salience of the target contrasts and robust phonological representations. Responses also highlighted the cognitive demands of processing multimodal input and orthographic incongruences, as well as environmental distractions. Higher performing participants appeared less affected by written input and better able to navigate cross-linguistic influences. However, it was evident that interference from incongruent L1 grapheme-phoneme correspondences persisted into high levels of proficiency. Taken together, these profiles moved beyond the previous operationalisation of proficiency to look at other individual differences relevant to understanding the influence of cross-scriptal orthographic input, such as learner awareness, L2 phonological development, cognitive control, language experience, and environmental setting.

8.2.5 Summary

The findings reported here provide evidence that orthographic influence depends on the perceptibility of the target contrast, as well as literacy experience in the scripts of the orthographic input. Literacy experience encompasses familiarity with a writing system, as well as experiences surrounding literacy development. For example, participant perceptions, beliefs, and strategies revealed varied ways that learners draw on how they have learned to learn in the past. Participants used existing knowledge of languages and literacies to create associations, analyse target forms, and encode new words in memory. However, potential gains associated with exposure to written forms are limited to familiar scripts, reliable grapheme-phoneme correspondences, and easy to perceive phonological contrasts. In other words, there is no evidence that any form of written input supported the encoding of confusable phonological contrasts. There was also no evidence that strategies drawing on written forms improved accuracy. It was only the metacognitive strategy of *evaluating learning* that marginally predicted accuracy, demonstrating that strategies may not relate to performance in the way that participants perceive them to.

The mismatch between the overwhelmingly positive perceptions of written input and the inhibited performance encoding difficult L2 contrasts with OI suggests two things. Firstly, learners are not well-positioned to assess the accuracy of the newly learned lexical representations. Secondly, they are often unaware of, or do not prioritise, lost attention to acoustic input, for a variety of reasons. The subsequent sections draw together overarching points of discussion which consistently emerged throughout each of the analysis chapters, in relation to key findings outlined above. Firstly, theoretical implications and contributions to the field are presented. This is followed by pedagogical applications of findings and the proposed theoretical position, including adaptations of materials for the classroom. Next, the methodological approach is discussed, in relation to the strengths and limitations of approaches to design and data collection, leading to recommendations for future research.

8.3 Theoretical contributions

As the evidence base for the complex and persistent influence of orthographic input on L2 phonological acquisition grows, the need for theoretical accounts of this influence becomes increasingly difficult to ignore. Additionally, the direct connection between this line of enquiry and the ubiquity of written input in instructed settings invites improved dissemination of findings to language teachers and learners, so as to increase awareness of factors relevant to language learning that may otherwise be overlooked. This section examines findings in relation to existing models of L2 speech learning and processing, and draws together preliminary theoretical contributions from the existing literature. The aim of which is to highlight areas that deserve further attention, as well as offer preliminary routes towards descriptive, explanatory, and predictive theoretical frameworks, which incorporate orthographic influence into L2 phonological acquisition and reach beyond shared scripts.

8.3.1 Building on models of L2 speech learning

At the start of this thesis (chapter 2.2.2), prominent models of L2 speech learning were outlined, namely PAM-L2 (Best, 1995; Best & Tyler, 2007; Tyler, 2019), SLM-r (Flege, 1995; Flege & Bohn, 2021) and L2LP (Escudero & Boersma, 2004; van Leussen & Escudero, 2015). These models were drawn upon to predict the L1 Arabic-speakers ability to perceive, recognise and produce words differing by the L2 English /f-v/ contrast. Particular emphasis was placed on the predictions of PAM-L2 and L2LP, as they focus on phonological contrasts. All three of these models predicted difficulty in relation to this contrast, based on (1) the perceptual similarity to the closest L1 sound /f/, and (2) the need to create a new category to accurately encode the contrast. This difficulty was born out in the findings. However, it was also the case that the L1 English group exhibited lower than expected performance with this contrast, indicating an issue of perceptual salience as well L1 perceptual bias. As perception and production tasks were not included in the final version of the study, it was not possible to examine the patterns of assimilation predicted by these models. Yet, their theoretical claims remain relevant to the discussion of findings.

With regards to L2LP, the argument that L2 speech learning is fundamentally meaning-driven aligns with points of discussion throughout the analysis. Many participants in both language groups reported the need to form meaningful associations between target items and known languages to support the items in memory. These associations often hinged on the initial sounds and letters, which were also the target contrasts, giving clues to meaning-based patterns of cross-linguistic influence. As noted in chapter 6.4.2.1, L1 Arabic open responses only mentioned lexical associations including /f/ but not /v/. This was not the case for the L1 English responses. Whether or not this is a replicable finding using more targeted testing remains to be seen. One route to further explore these ideas builds on L2LP references to error-driven learning, which is discussed further in section 8.3.2.

As mentioned in chapter 2.2.2, Tyler (2019) extended the predictions of PAM-L2 to language classrooms, paying specific attention to written input. Comments centred on the influence of orthography in relation to perceptual equivalence classification, as well as the relationship between written materials and rapid L2 vocabulary expansion. Regarding differences that are easier for an L2 learner to perceive, Tyler predicts that alphabetic orthographies which clearly (or congruently) signal a phonological difference may help to tune into phonetic differences in the speech signal. However, when the orthographic signal is not clear "their internal rehearsal of the pronunciation of L2 words via orthography may reinforce a perception that the L2 phonemes are equivalent rather than distinct" (Tyler, 2019, p. 617). While not explicitly tested, these claims broadly align with the findings of the present study in several ways. Firstly, the importance of written input for rapid word learning was repeated in responses, aligning with other studies that argue written over audio forms can facilitate more accurate recognition of novel minimal pair words (Escudero et al., 2022)[49]. Future research would benefit from a more systematic investigation of rapid word learning with written input and the potentially detrimental effects on establishing phonological representations of confusable contrasts. Secondly, while there is a lack of detail in what Tyler means by a "clear" orthographic signal, these ideas are supported by empirical findings, in the present study and existing literature (Hayes-Harb & Barrios, 2021), with regards to (in)congruence and perceptibility. Specifically, shared-congruent OI can facilitate the acquisition of contrasts, when a phonological contrast is perceptually salient. However, shared-incongruent OI is likely to interfere with lexical encoding of contrasts, whether they are perceptually confusable or not.

It is interesting that Tyler mentions the rehearsal of pronunciation via orthography, as this was another recurrent theme within the findings of this study. L1 Arabic participants reported the usefulness of orthography as a guide for pronunciation, as well as frequent reference to visualising the spelling of lexical items and initial sounds, whether or not they had been exposed to the spelling. This was again highlighted in the PCA analysis of the POLLS inventory responses, where *producing* strategies brought together the items *repeating* and *mouthing* of words with *using English spelling to distinguish similar sounds*. Thus, indicating the influence of orthography on both perception and production, and the connection between the two modalities.

Overall, the findings from the present study confirmed the anticipated difficulty of lexically encoding the /f-v/ contrast, as an example of a phonological contrast that is not well-established in L1 phonology and easily assimilated to a single L1 category /f/, predicted by all three models. Open responses also underscored that learning was meaning-driven, as well as frequent indications that participants were focused on learning the phonological

[49] In this study, participants were exposed to either written or audio input, not in combination.

forms of the target items, both of which were reportedly supported by orthographic knowledge. These findings support the attention given to lexically-driven L2 speech learning across all three models, but further challenges the absence of orthographic influence in formalised accounts. The speculative predictions of PAM-L2 in relation to written input are broadly supported by the present study. However, there is scope for predictions to be more directly tested and a need for inclusion of cross-scriptal orthographic influence. While an emphasis has been placed on PAM-L2, being the only model to include explicit references to orthography, the computational approach proposed in L2LP potentially offers a better option for modelling orthographic influence, based on its interest in the whole developmental trajectory, connection to error-driven learning, and ability to capture perceptual biases and constraints. This is discussed in more detail in the next section.

8.3.2 Alternative perspectives on orthographic influence

Continuing to think about orthography and the relationship between perception and production, Nagle and Baese-Berk discuss a dynamic perspective of how that relationship may change over the course of learning, noting that:

> Once perception begins to stabilize, entering a developmental plateau, the perception-production link itself might also begin to stabilize, such that no change is observed in the link during, or the effect of perception on production may decrease in strength or disappear altogether. (2022, p. 585)

The stabilising of perception at a point of developmental plateau, and the subsequent influence on production may connect to Tyler's (2019) predictions related to written language materials, rapidly increasing vocabulary development, and fossilisation of perceptual learning. Could Tyler's predictions be interpreted to mean that exposure to orthographic input and reliance on written language materials can prematurely stabilise L2 perception? In order to understand the learning mechanisms at play here, an error-driven, discriminative perspective could prove insightful. For example, the aim of error-driven learning is to reduce uncertainty through assessing the difference between expectations and outcomes, then updating future expectations based on experience. In other words:

> Error-driven learning tries to achieve optimal discrimination of cue structures by minimising the error between the desired state of full certainty about an outcome

> and the actual current expectation of this outcome to occur given the cues that are present at that point in time (Hoppe et al., 2022, p. 2227)

The weighting of cues is then adjusted based on both frequency and how predictive/useful the cue is for an outcome. When multiple cues relate to an outcome, as is the case with multimodal input in the present study, there is competition between them and blocking of certain cues, depending on the assessed informative value of each.[50] While these ideas overlap with associative learning constructs (N. C. Ellis, 2006, 2008), a discriminative perspective emphasises the role of prediction over association.

Qualitative and quantitative analysis of participant perspectives on their learning highlighted frequent mention of feeling more confident or certain about word forms when written input was available, as well as considering written forms to be "correct". This provides evidence that OI reduced uncertainty about the representation of the target forms, potentially blocking competing acoustic cues. The premature reduction of uncertainty based on the precision offered by written forms may then prevent the continued updating of expectations or, in other words, the fossilisation of representations. This account goes some way towards explaining the disunity between participant performance and perceptions about their learning, particularly in relation to a reliance on orthographic cues at the expense of sufficient phonetic detail in L2 phonological representations. Further exploration of orthographic influence with computational approaches, such as error-driven learning, would be a valuable direction for future research.

These ideas align with work by Veivo and colleagues (2015, 2016, 2018; 2013), who have explored orthographic bias in adult L2 lexical knowledge and imprecise phonological representations. Specifically, they posit that late L2 instructed learners' orthographic representations may be more robust than phonological representations, although this bias may decrease with proficiency. Cutler (2015) also reiterates the point that without a firm perceptual basis, a lexically encoded distinction via orthography does not avoid L2 confusion when perceiving and processing spoken language. Knowing the distinction exists is not enough to perceive it correctly in speech.

[50] Outcomes can also compete with each other, with connections made to asymmetric effects based on cue and outcome competition ratios (Hoppe et al., 2022, p. 2227).

To explain their findings, Veivo and Jarvikivi (2013) draw on activation accounts of L1 spoken word processing to argue for a co-structuration account. The authors build from online co-activation accounts, where orthography and phonology are linked but separate codes, and the restructuration account, where orthographic information changes existing phonological representations into amalgamated representations of both orthographic and phonological information. They then go on to suggest that, during instructed L2 learning:

> If orthographic forms of words are learned first or simultaneously with the corresponding phonological forms, we could hypothesize that lexical representations so formed would contain both orthographic and phonological information from the early stages of L2 learning. (Veivo & Jarvikivi, 2013, p. 866)

These ideas are then applied to improving levels of proficiency, assuming that lexical representations would increase in both stability and accuracy of phonological and orthographic information. This account may shed light on the conflation in participant responses of sounds and letters, indicating the difficulty separating phonological and orthographic information in the mind is because they are not separate. These views then echo claims from the lexical quality hypothesis, where underspecified phonological forms lead to increased competition, while improving precision, for example with increasing proficiency, limits interference (Cook et al., 2016; Perfetti, 2017; Perfetti & Hart, 2002).

Another perspective that could be of interest is that of *desirable difficulty (R. A. Bjork & Bjork, 2020; Bogulski et al., 2019; Suzuki et al., 2019)*, mentioned in chapter 5. For example, "desirable difficulties are conditions of learning that impose initial costs by inducing errors, requiring conceptual elaboration, or increasing the requirement to negotiate variation, but benefit learning and memory over time" (Bogulski et al., 2019, p. 1052). Research has found that, with increased proficiency, bilinguals become more proficient at regulating their L1 and inhibitory control, echoed in the findings of the higher proficiency learners in the present study. Looking at individual differences, some studies have found that high-skill learners benefit more from desirable difficulty than low-skilled learners, such as word learning by high vs. low skilled spellers (Eskenazi & Nix, 2021) and word learning through dominant vs. non-dominant language translations (Bogulski et al., 2019). The fact that participants relied on orthographic input to reduce the difficulty of the task would be interesting to explore in contrast to the potential of audio input offering more *desirable difficulty*, with longer-term benefits for new words in memory. However, this is then likely to interact with individual differences, such as proficiency and inhibitory control. Alternatively, incongruent multimodal input may be an example of undesirable difficulty.

Finally, participant perceptions about the importance of written input for supporting the storage of forms in memory, speed of learning, and references to internal rehearsal of both audio and visual input highlights the relevance of working memory (WM) for research in this area. Considerable research has investigated individual differences related to WM and its connection to L1 and L2 acquisition and processing, including specific reference to learning the phonological and orthographic forms of vocabulary items (Baddeley, 2017; Koda, 2007; Quam et al., 2018; Z. E. Wen et al., 2020; Z. Wen & Li, 2019). Of the different components of WM, two have received the most attention in relation to language acquisition, namely phonological short-term memory (a.k.a phonological loop) and the attention-oriented "central executive" component (Z. Wen & Li, 2019). The visuo-spatial sketchpad and episodic buffer components, which relate to visual/spatial information and the integration of information from different modalities to long-term memory (Baddeley, 2012, 2017), are also of clear relevance to the present study. Measures of WM are rarely reported in studies investigating orthographic input on L2 phonological acquisition, yet this is likely to be a fruitful line of enquiry, alongside other individual differences.

8.3.3 Predicting orthographic influence across writing systems

Drawing together these different perspectives, there appears a convincing and coherent argument regarding the perceptual and processing cost of early reliance on orthographic input and broader written language materials. As has been discussed, there are scenarios where orthography can support learning, such as speed of vocabulary acquisition and enhancement of categories that are perceptually salient, when the grapheme-phoneme correspondence is congruent between both languages. However, there are many scenarios where L2 learners are faced with the task of learning perceptually difficult sounds, with added layers of orthographic incongruence and script familiarity or distinction to navigate. Therefore, laying the foundational building blocks of L2 speech and vocabulary learning with both phonological and orthographic information may result in persistent orthographic bias and interference into high levels of L2 proficiency.

To visualise predictions based on the findings and discussion presented in this chapter, figure 8.2 offers a preliminary hierarchical model denoting scenarios of orthographic facilitation and interference with regards to L2 phonological encoding. It starts from the assumption that perception is the basis of L2 speech development and spoken language processing involves the activation of phonological information in mental lexical representations. This model integrates ideas from L2LP and co-structuration accounts, which also relate to proficiency. For example, it is important to account for the fact that the quality of L2 phonological representations may change with proficiency, which would then affect the predicted influence of orthographic information. In general, if a learner does not have a robust representation of target phonology, then orthography is not predicted to support L2 phonological distinctions (Escudero, 2015).

The next layer distinguishes between systematicity and congruence, based on proposals by Hayes-Harb and Barrios (2021). Orthography is predicted to interfere with L2 phonological acquisition when GPCs are not transparently reliable[51] (systematic one-to-one mappings), even when target phonology is well-established for the learner (Hayes-Harb et al., 2010). A limitation to consider is that this does not capture L1-related expectations of transparency (Erdener & Burnham, 2005), which qualitative analysis in this study also pointed towards. For example, L1 English participants made more mention of spelling not reflecting pronunciation, which may reflect L1 literacy in an opaque orthography. In contrast, consonant sound and letter mappings are highly reliable in Arabic written forms, although orthographic depth is more debatable (see chapter 2.3.1). Finally, this model predicts that even if target phonology is well-established and L2 GPCs are transparent, they may not be congruent with the GPCs of the L1, again introducing interference rather than facilitation effects (Bassetti et al., 2018). This illustrates the argument that early reliance on L2 written language is, at most, minimally facilitative of precision in developing L2 phonology. Meanwhile, there is ample opportunity for orthographic interference, which may reduce with increasingly robust phonological representations but is likely to be a persistent influence, despite proficiency gains.

[51] The terms transparent and reliable are used to reflect the wording of participant responses.

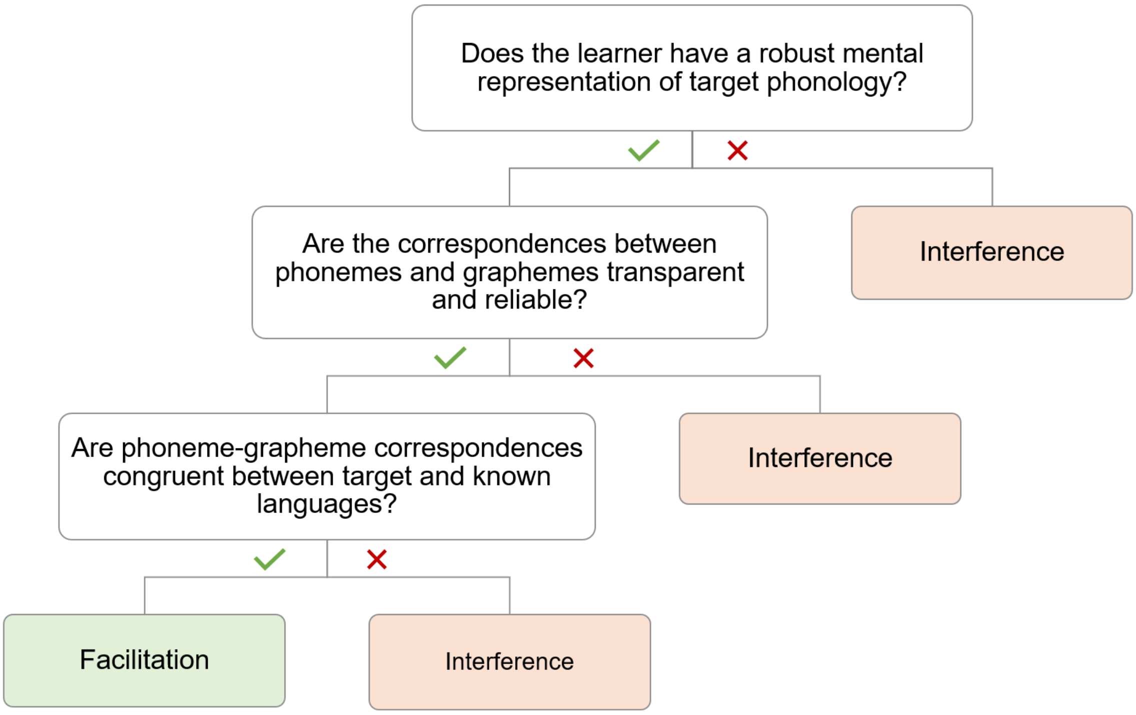


Figure 8.2 Predictions of minimal facilitation of orthography for L2 phonology

Extending the illustration in figure 8.2 to the findings in this and other cross-scriptal studies, figure 8.3 provides a diagram of predicted orthographic influence including factors that relate to learning languages across writing systems. The illustration starts from the same point but introduces script distinction, offering two divergent paths depending on whether the L1 and L2 share the same script. If the scripts are shared, then the same predictions for figure 8.2 apply. If the scripts are not shared, then the next question to consider is whether learners are sufficiently familiar with the orthography to make use of the information presented. If the script is entirely unfamiliar and encoded phonological information is inaccessible, then minimal interference is anticipated, due to visual disorientation rather than cross-linguistic influence, in this case (Mathieu, 2016; Showalter & Hayes-Harb, 2015).

If learners are sufficiently literate in the L2 script, the next question concerns the transparent reliability of the sound-graph correspondences. This is similar to the second level of figure 8.2; however, there are two important distinctions. Firstly, sound-graph correspondences are used to be inclusive of diverse grainsizes of writing systems (alphabet, abjad, syllabary, logographic etc.), rather than phoneme-grapheme relationships that only relate to alphabets. Additionally, there is no mention of congruence, as this is only relevant when L1 and L2 orthographies share the same script. If sound-graph correspondences are not reliable, then this introduces interference. Thus, the only scenario where cross-scriptal orthographic input facilitates L2 phonological development is when (1) the learner has a precise mental representation of target phonology, (2) is sufficiently familiar with the L2 script, and (3) the L2 phonology is reliably represented in the orthography.[52]

With regards to the present study, this model accounts for the fact that lexical encoding accuracy for /m-n/ contrast words was facilitated by both Arabic and English OI. For Arabic participants, this is a well-established contrast for learners, they are literate in both orthographies, and in both cases sound-graph correspondences are reliable. Additionally, in the case of Arabic orthographic input, the GPCs were congruent between the L1 and L2. Meanwhile, as the /f-v/ contrast was likely represented with varying degrees of precision, related to English experience and exposure, this correctly predicts broad orthographic interference, regardless of script input. For high proficiency learners, the lack of reliable orthographic information regarding the distinction between /f-v/ in Arabic OI results in persistent interference effects. However, as the mental representation of the /f-v/ contrast increases in precision, learners are better able to integrate the distinct, familiar and reliable orthographic information from the English written input, to support the accurate encoding of the target contrast.

[52] The binary divisions reflected in this model are unfortunate but pragmatic. It is acknowledged that proficiency, literacy, orthographic transparency, script overlap and amount of facilitation or interference are not dichotomous, but vary along a scale.

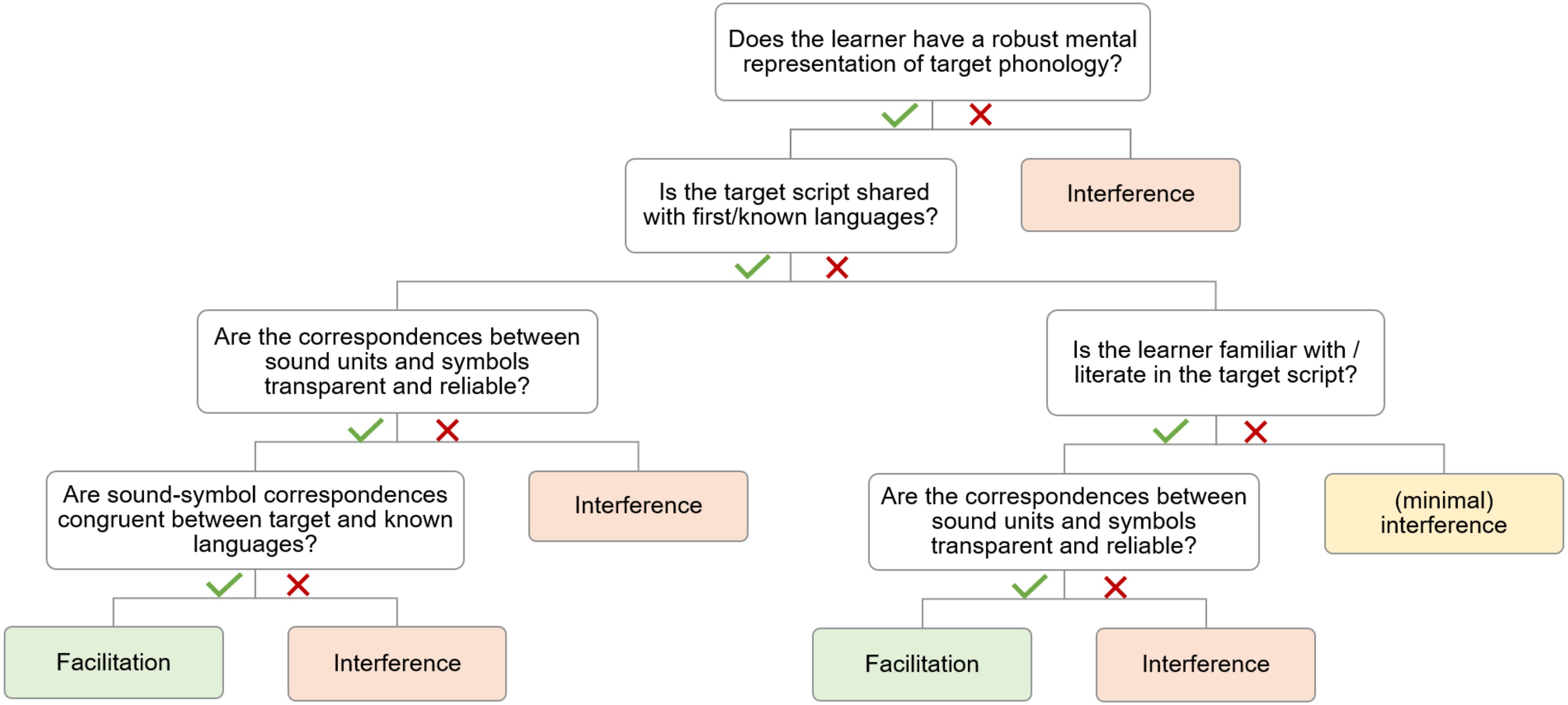


Figure 8.3: Predictions of minimal facilitation of cross-scriptal orthography for L2 phonology

In summary, this section has discussed relevant theoretical accounts of L2 speech and spoken language processing, in order to construct a preliminary account of orthographic influence on L2 phonological acquisition. The predictions in figure 8.3 demonstrate how the original research question, regarding whether different script input has a varying influence on lexically encoding a confusable contrast, can be answered with both yes and no. This model captures learner perceptions that different script input has a distinct influence on word learning, as well as the variety of ways in which orthographic information can lead to interference or facilitation across writing systems. This accounts for the broadly inhibitory influence of orthographic input accompanying /f-v/ contrast words and the facilitative effect with /m-n/ words, by centering the importance of perceptual difficulty and precision of phonological representations.

The model has the flexibility to incorporate the influence of proficiency, where phonological representations are likely to improve in precision with increased experience. Further, the influence of written input is likely to change depending on literacy in the target language orthography. This would benefit from further individual difference testing in future research, particularly in relation to literacy skills. This model currently assumes literacy in the L1 orthography, which does not represent all adult L2 language learning. Additionally, it remains unclear at what point, when developing literacy in a new writing system, orthographic information is sufficiently intelligible to be integrated with lexical representations. Finally, the account presented here captures the variables of systematicity, familiarity and congruence, but frames these in a way that more closely resembles participants' articulations around concepts of script distinction, familiarity and reliability.

8.4 Pedagogical applications

The previous section presented an argument that early exposure to orthographic input only facilitates L2 phonological acquisition in a limited set of circumstances, fundamentally dependent on the ability to perceive target L2 phonology. This perspective aligns with broader recommendations against an overreliance on early orthographic input, or at least some pedagogical caution, in instructed environments (Hayes-Harb et al., 2018; Rafat & Stevenson, 2018; Sokolović-Perović et al., 2019; Tyler, 2019). Discussion of pedagogical implications within this field of research is limited, and there is a known disconnect between

L2 phonology research and pronunciation instruction more broadly (Darcy, 2017). In their review, Hayes-Harb and Barrios (2021) highlight that the two main orthographic effects relevant to teaching are the potential for (1) systematic orthographic input to facilitate the encoding of target phonology, and (2) incongruent grapheme-phoneme correspondences to cause interference. The opportunity for both facilitation and interference relate directly to the discussion above, specifically predictions in figure 8.3. In fact, this visualisation of orthographic influence was created with a pedagogical, as well as theoretical, intention.

One aim of visually demonstrating that there may be more potential for interference than expected, is to persuade teachers that this is something worth considering as part of their practice. My anecdotal experience as a language learner and teacher is that people take some convincing about the drawbacks of relying heavily on written input during language learning, and there is a lack of clarity around how to avoid early exposure to orthography. With this in mind, a "teacher-friendly" version of figure 8.3 was created (appendix XXIII), which could prove a useful device for teachers to consider the best way to either introduce new vocabulary or draw learners' attention to points of anticipated interference.

Additional concrete suggestions have been offered by Rafat and Perry (2019), focused on L1 English learners of L2 Spanish. They suggest an exercise introducing auditory input before exposure to written forms, in order to mitigate potential interference. Another lesson plan recommends targeting the trill and flap rhotic contrasts, intentionally making use of facilitative systematic spellings i.e. /ɾ/ - <r> and /r/ - <rr>. Tyler (2019) also made several suggestions for ways to incorporate the principles of PAM-L2 into language classrooms, particularly focusing on early perceptual training of phonetic differences, before substantial L2 vocabulary acquisition. For example, he recommends delaying the introduction of orthography, both to increase the opportunity to learn phonetic differences and to delay the rapid expansion of L2 vocabulary. Alternatively, he suggests teaching IPA at the beginner stage and later introducing L2 orthography. Other suggestions include the use of high variability perceptual training, the drawing of attention to L2 phonological contrasts in pronunciation, or the inclusion of perceptual assimilation tests as part of student diagnostic assessments.

With regards to teaching vocabulary, Tyler recommends taking perceptual assimilation into account, such as starting with easily discriminable categories and introducing single-category contrasts more slowly, accompanied by perceptual training activities. The idea of delaying exposure to orthography and instead drawing attention to L2 phonological contrasts and articulatory gestures influenced the creation of the videos shared with participants after completing the study (chapter 3.3.1). These videos modelled and contrasted English consonants and vowels, with two videos dedicated to sounds that are likely to be difficult for L1 Arabic-speakers, in particular. Additionally, these videos were created with teachers in mind, who could model a similar approach in their classroom.

While there may be benefits of prioritising high quality auditory input to facilitate early perceptual learning, the avoidance of orthographic input in the classroom is "both unrealistic and in all probability too simplistic"(Escudero et al., 2014, p. 394). This is reflected in the results presented in chapter 5, where almost 90% of the L1 Arabic group stated that it was important to see the spelling when learning new words in a language. The emphasis on consolidation, memorisation and speed of learning in participant responses, and the role of written language in the rapid expansion of L2 vocabulary (Tyler, 2019), is highly compelling for language learners and teachers. Furthermore, the few studies exploring interventions to overcome orthographic interference have yet to report improvement as a result of explicit instruction (K. Brown, 2015; Hayes-Harb et al., 2018; Showalter, 2020).

These patterns also resonate with my personal experience as a language learner and teacher, where I have experimented with avoiding early exposure to written forms, particularly when presenting new vocabulary and focusing on pronunciation. For example, as a teacher, my students have demanded to see written forms, often with frustration and disorientation, when learning new vocabulary. Equally, as a learner, when asking a teacher for more auditory exposure rather than written language, I have been met with confusion and sometimes hostility. That being said, when learning Arabic at a language school which used the Growing Participation Approach (GPA)[53] - which involves conversational, immersive, structured play in a language - I, too, felt regularly frustrated and disorientated that all initial communication was orally-based. Also, in my early Arabic classes, before being able to read the Arabic script, I found myself quickly resorting to transliterations and personal codes in my notes, to support and anchor my learning. All of this to say, as is well-established, adult language learners bring all their knowledge, habituated processing of

[53] https://www.growingparticipation.com/our-approach

known languages, and life experiences to the learning of additional languages. Of course, literacy experience is often part of the package. The question is then, how can we equip teachers and learners to make the best use of the knowledge and resources available to them, in light of empirical research into orthographic influence on L2 phonology?

The analysis of participant language learning strategies in chapter 6 holds some possible routes to explore, particularly in relation to the development of metacognitive awareness and strategies. Of the four key strategy components, *producing new words*, *analysing sounds and letters*, and *using L1 Arabic OI for support* did not predict improved performance, whereas the metacognitive strategy *evaluating learning* was the only component that offered some relationship to improved accuracy. This suggests perhaps the best way to support learners is to encourage the development of metacognitive strategies, where learners are not directed towards one approach, but rather increase awareness of different influences on their learning. One way to do this could be through the classroom application of the POLLS inventory developed for this study.

As mentioned in the research context provided in chapter 6, the POLLS inventory, like other LLS inventories, was inspired by the SILL (Oxford, 1990). The SILL, while a controversial tool, has been conducted and adapted for teaching and research all over the world, remaining the most frequently used research instrument in LLS research. Amerstofer (2018) draws particular attention to the original purpose of the SILL, as a self-evaluation tool for L2 learners, and the success of its user-friendly and comprehensible design. It is with this in mind that the present study advocates for use of the POLLS inventory to encourage metalinguistic awareness and evaluation of strategy usage in pedagogical settings. A modified version of the POLLS inventory is provided in appendix XIX, including an introduction and suggested lesson integration, aimed at teachers. This tool is not intended to promote a specific approach to teaching or learning, but rather offers discussion-based activities, reflective practice, and autonomy for learners to become more aware of what influences and benefits their learning, in light of their personal learning goals. Additionally, the statements themselves may introduce ideas that learners may not have been previously aware of and inspire future experimentation with, and evaluation of, LLS usage.

The inventory has been amended and expanded in response to limitations discussed in chapter 6, such as extending the inventory to 50 items (the same number as SILL). This extension will improve the quality of analysis, if used in a research context, as the issues that arose from a lack of correlation between items and low sampling adequacy should be reduced. Furthermore, the extended version incorporates additional strategies mentioned in the qualitative analysis and revisits relevant items from existing strategy inventories (Berkil, 2008; Oxford, 1990; Pawlak, 2010; 1997; Rokoszewska, 2012). This includes adapting items from the *Strategy Inventory for Character Learning (Shen, 2005)*, which were absent in the original version but prove useful in explicitly targeting orthographic knowledge. Section A of the modified POLLS inventory begins with four open questions, which lend themselves to introductory discussion. This is followed by 50 Likert-scale items, which are grouped in relation to the strategy components extracted during the PCA in chapter 6.5.2, as follows:

- *B.1*: How learners initially learn new words
- *B.2*: Use of L1 orthographic knowledge
- *B.3*: Associating and analysing words, sound and letters
- *B.4*: Practising pronunciation
- *B.5*: Metacognitive strategies

Therefore, the modified version of POLLS is well-positioned for pedagogical application and further research, where it could be translated or adapted for other language settings.

A final pedagogical point to consider is the focus of the present study on intermediate and advanced proficiency students. The study was originally designed with beginner learners in mind, particularly those with limited literacy or formal schooling in their first language(s). Just as the online version of the study proved inappropriate and inaccessible to these learners (Shepperd, 2022), the same can be said of the classroom application of the POLLS inventory. However, the broader findings of the present study continue to have relevance for learners at every level, and highlight the importance of further research and pedagogical advancement, in relation to orthographic input and phonological development. For example, the discussion of proficiency-related variables demonstrated that high proficiency learners may be better able to mitigate negative effects of orthographic input and take advantage of facilitative information, yet orthographic inference persists. The persistence of orthographic influence and interference, despite advanced proficiency and cross-scriptal literacy, underscores the importance of considering the timing and amount of exposure to written language during the earlier stages of language learning.

This is even more stark when considering learners who are not only learning literacy in a new script, but who are learning literacy for the first time in a language that they are not familiar with. Of course, there is a pragmatic urgency to equipping language learners with literacy skills, particularly in the contexts of migration and social integration. However, the consequences of early exposure to simultaneous written and spoken language, and the possibility of alternative approaches for this population of learners, evidently deserves greater attention.

8.5 Methodological considerations

There were a number of novel aspects to the research design used in the present study, including the prioritisation of participants' perspective within a psycholinguistic design, the use of internet-based methods, and the development of a novel LLS inventory tool, all in the context of L2 orthography and phonology research. This research also builds on and extends studies investigating orthographic influence across writing systems. To date, orthographic influence has been investigated in relation to Arabic by looking at naïve L1 English learners of Arabic pseudowords (Jackson, 2016; Mathieu, 2016; Showalter & Hayes-Harb, 2015), beginner to intermediate L1 Arabic learners of L2 Chinese (Zhang & Roberts, 2019, 2021), and beginner L1 Arabic learners of L2 English, including adults with limited L1 literacy (Al Azmi, 2019). This highlights a lack of research with higher proficiency language learners, which Plonsky (2023) notes is a pattern across applied linguistics research, as more studies recruit samples of beginner and intermediate learners. Additionally, of the studies mentioned in relation to Arabic, neither of the two studies recruiting L1 Arabic-speakers focus on the acquisition of segmental phonology in relation to orthography. The present study offers evidence in relation to both under-researched areas. However, the aim to design a multi-method psycholinguistic study, which was also accessible to adult migrant language learners with varied L1 literacy experience, was not a straightforward process. Thus, design choices and their consequences are further discussed here, with particular focus on online adaptations, recruitment, stimuli design, and data quality.

Due to the need to rapidly respond to national lockdowns related to the Covid-19 pandemic, the study was redesigned multiple times to run online. This led to changes which ultimately strengthened the project, but also others which would need to be addressed in future research. To reduce the risk of high dropout rates and technical difficulties, the experiment was designed to last around 30 minutes and any elements that increased difficulty for participants or varied in data quality were removed from the design. This resulted in the reduction of target contrasts and stimuli items, the removal of eye-tracking, perception, and production tasks, as well as the battery of individual difference measures (phonological awareness, working memory test, vocabulary size test). The importance of including proficiency and ID measures was reiterated in chapter 7. As such, future research building on this design would benefit from more robust proficiency measures, such as a vocabulary size test, and working memory tests as highlighted in section 8.3.2. Perception and production tasks would then facilitate investigation into the specific predictions of L2 speech learning models, in relation to orthographic influence. Finally, eye-tracking adaptations would be useful for further investigation of participant awareness, attention, and online processing, in relation to different input exposure.

Overall, the decisions to streamline the present design were well-motivated, resulting in a larger and more diverse[54] sample than would have been possible in-person, as well as the creation of a skeleton design that could be adapted/extended to explore several different avenues of investigation. Additionally, the decision to prioritise participant reflections through the post-test questionnaire proved a novel and rich perspective. It is also worth noting that an extended version of the experiment was built and piloted in Gorilla (chapter 3.2), which is ready to be implemented in a future study. This is of particular interest now that in-person and hybrid data collection is an option, removing some of the previous barriers to participation and issues around data quality.

As noted in chapter 2.2.3, studies investigating orthographic influence on L2 phonology often rely on relatively small sample sizes (~30). A strength of the research presented here is the far larger sample size (~115 in each language group), facilitated by both a within-subject design and the use of internet-based methods. However, there were several biases in the sampling, such as the high proportion of L1 Arabic participants who were (1) advanced learners of English, (2) Saudi Arabian, (3) female, and/or (4) had a background in

[54] As discussed in Shepperd (2022), inclusion of certain language learners was not possible. However, the design was inclusive to a wide range of participants who would not have been able to attend an in-person study, due to geography, availability, or apprehension. People participated from all over the world, with a range of devices, proficiency levels and educational experiences.

(English) linguistics. This is a disadvantage of online convenience and snowball sampling methods, accompanied by the discovery that email was not the best mode of communication with L1 Arabic participants, in many cases. Future research would benefit from exploring ways to connect participants to the experiment through social media platforms, such as Facebook, as well as in-person versions. The use of platforms, such as Gorilla hold exciting promise for taking psycholinguistic studies out of the lab, even if fully internet-based approaches are not appropriate. For example, when seeking to include underrepresented and potentially vulnerable language learners, it is important to conduct research in safe, familiar environments and remove geographical and financial barriers around travel. The present study could easily be conducted using laptops, tablets, or phones after language classes, with the researcher and/or teacher present. This would help address the need for additional support and encouragement during participation, as well as providing the opportunity for participants to ask questions and ensure study information is clear. In fact, it could be seen as an advantage that the study has been tried and tested with this sample of language learners before being extended to more complex learning contexts, with additional layers of vulnerability to consider. Another unforeseen advantage was the inclusion of a greater range of proficiency levels as discussed in chapter 7.

With regards to stimuli design, some questions were raised around the difficulty of learning pseudowords that are accompanied by novel images with low nameability. It is possible that these images increase the artificiality of the task, influencing the strategies used to remember the words and associated images. Indeed, participants often described strategies to form meaningful connections so as to "make sense" of the images. I do not view this as problematic, as it remains consistent with theories around input processing. For example, "learners process input for meaning before they process it for form, implicating a meaning-based processing as the default processing approach"(Han & Liu, 2013, p. 150), derived from input processing principles developed by VanPatten (1996, 2004, 2015). These ideas also fall within broader generalisations of functionalist/associative approaches (N. C. Ellis, 2006, 2008) and the meaning-driven learning advocated in L2LP (van Leussen & Escudero, 2015). There may be additional difficulty associated with the multiple levels of novelty. However, participant reflections imply that sound-picture referential meanings formed during this study mimic those found in natural language contexts and reported in associative learning research. Furthermore, I would argue that the ecological validity and inclusivity of the colourful and realistic stimuli used in the present study is an improvement on the use of black and white line drawings of abstract (Escudero et al., 2008; Hayes-Harb & Hacking, 2015) or real objects (Mathieu, 2016; Showalter, 2018; Showalter & Hayes-Harb, 2015). Participant engagement with the present study has also solidified my belief that the use of novel objects reduces risks around inadvertent deception and misuse of pseudowords in real world contexts.

Another issue raised with the stimuli was the perceptual difficulty associated with the /f-v/ contrast, for both L1 English and L1 Arabic participants. Closer analysis of the audio stimuli revealed evidence of partial devoicing of the words beginning with the voiced labiodental fricative /v/, which could explain the reduced salience of the contrast. Even with the added difficulty, the L1 English control group were still able to perform above chance when it came to lexically encoding the /f-v/ contrast, further supported by responses to the post-test questionnaire. Thus, this difficulty does not undermine the validity of the findings. However, interpretation of these findings was hindered by the limited number of stimuli and trials. Going forward, it would be of interest to adapt the study with an increased number of trials and stimuli per condition and/or investigate a different confusable contrast, such as /b-p/, to see if the findings are replicated. The extended version of the study was designed with another contrast in mind, namely /θ-ð/, which is systematically represented in Arabic but not English. This was in order to test whether unreliable English distinct script input introduced interference when a contrast was established in both the L1 and L2.

Other considerations for future applications of this study design include the advantages of within and between subject designs for stimuli presentation. For example, participants noted the additional processing demands of switching between different script inputs during word learning. To reduce such switching costs, each block of trials during word learning only contained one form of written input. However, it was necessary to switch orthography between learning blocks. This is certainly a more artificial aspect of the experimental design, and may have influenced participant reflections and preferences in the post-test questionnaire. Relatedly, due to the randomisation of trials, some participants were not exposed to any written forms until midway through the learning phase, which makes it difficult to interpret responses to Q1 about when participants first noticed the written input, or whether they noticed it at all. Observations during piloting and responses to the post-test questionnaire make it clear that participant awareness of, and attention to, written input varied, which deserves further investigation. This line of enquiry could explore theoretical understandings of noticing, attention, awareness, and input processing (Robinson et al., 2013; Schmidt, 1990; VanPatten, 2004), in relation to written input and L2 phonology.

Continuing to think about insights into awareness and attention, a strength of the present study is the use of mixed methods approaches and the incorporation of participant perspectives, alongside behavioural data. Some limitations of using written questionnaires included reduced data quality, where participants provided very short or unclear answers, as well as misunderstanding the questions. For example, confusion around Q6 (*Did it make a difference seeing the written words? In what way?*) by the L1 Arabic group led to a large number of excluded responses. Additionally, in both groups some participants misunderstood that Q8 (*In general, when learning new words do you think it is important to see the spelling the first time you hear it? **Why?***) related to broader learning experiences rather than the experimental task. The extended design made use of interviews to facilitate better understanding of participant perspectives, and I believe this would be fruitful to explore going forward. In relation to the present study, questionnaires were an appropriate choice, in terms of time and resources. There was also the advantage of participants recording their responses immediately after completing the study, avoiding the delay of organising an interview (Robinson et al., 2013).

Another factor, when considering optimal conditions for this type of study, was level of distractions. This is particularly difficult to control for in internet-based studies; however, self-reported information into amount and type of distraction proved insightful, where participants were surprisingly honest and detailed in some cases[55]. Most participants reported little to no distractions, but some frequent examples included interruptions from a household member, tiredness, or background noise. The L1 Arabic group reported higher levels of distraction, on average, and this proved to be a significant predictor of accuracy in the matching task. This demonstrates the usefulness of collecting distraction data, both for informing exclusion procedures and to include as part of analysis. Alongside the distraction check, the use of audio checks, attention checks, and read aloud text proved useful additions

[55] As mentioned in chapter 3, some participants reported alcohol consumption, multi-tasking while eating dinner, and specific situations which caused unforeseen interruptions.

to the study design. In particular, the read aloud function, which was designed to improve accessibility for participants with limited literacy, resulted in participants generally reporting better engagement with instructions and improved experience.

A final point in relation to methodological choices and interpretation of findings relates to the move from a multi-session design to a single-session experiment. Arguably, single-session word learning studies offer limited insight into learning and reflect processes more associated with short-term memory (Bakker et al., 2014), especially without a consolidation period of sleep (Dumay & Gaskell, 2012, 2007; Gaskell et al., 2014). Others argue that lexicalisation of newly learned words can take place rapidly and without consolidation (Kapnoula & McMurray, 2016), and there is clear precedent in the field to draw conclusions based on single-session studies. Even so, to gain more insight into learning trajectories over time, more research should move beyond single-sessions. The final study did include a delayed post-test after 2 days; however, the low take-up (~35 participants from each language group) led to the decision not to analyse this data as part of this thesis. The low take-up is likely related to the use of internet-based methods and lack of sufficient incentive to return to the study. This could be addressed through more consideration of appropriate incentivisation and, as mentioned before, the incorporation of in-person data collection, where you can take the study to participants after set time intervals. Future analysis of the dataset from the delayed post-test provides the opportunity for preliminary insights that can guide future multi-session designs, with the added advantage that the extended four session experiment has been built and piloted in Gorilla already.

Overall, methodological considerations discussed above highlight the strengths and limitations of the present study. Looking forward, the richness of the findings, despite contextual challenges around the shift to online data collection, provides even stronger motivation for pursuing this line of enquiry, and extending the design and instruments developed here. Additionally, in the short time since data collection for this study took place, technological developments have already addressed many challenges around internet-based eye-tracking and audio recordings that prevented their inclusion in the present study. Thus, these elements would be easier to incorporate in future research, as well as offering more reliable data into perception, production, and processing.

8.6 Summary of future directions

Throughout this discussion several avenues for further research have been highlighted. In the theoretical contributions, connections were made to error-driven and computational approaches to understanding learning mechanisms and the processing of multimodal input. The value placed on written input and the role of orthography in reducing uncertainty, with the potential to block attention to acoustic cues and attunement of lexical representations, is of particular interest. Additionally, a model with testable predictions was posited to further investigate cross-scriptal orthographic influence on L2 phonology. Specifically, this account offers the opportunity to develop predictive weightings and insights into probabilities of facilitation and interference. Furthermore, as highlighted in both the theoretical and methodological discussion, an extended version of this study with perception and production testing would prove useful to investigate how orthographic influence relates to the development of robust mental representations of confusable L2 contrasts. This would also address continuing queries around the relationship between proficiency, L2 phonological development and orthographic input. Proficiency can seem a nebulous concept, thus it would be useful to investigate relevant individual differences, including measures of language experience and exposure, cognitive capacity and language learning strategies, beliefs and goals.

With regards to pedagogical directions, a modified version of the POLLS inventory was proposed for use in both research and classroom settings. The extended number of items, including insights from the qualitative analysis and relevant existing inventories, is well-positioned for future testing and validation as a measure of phonological and orthographic LLS usage. Additionally, a classroom application was suggested, with the aim of developing increased awareness, in both teachers and students, of different pronunciation learning strategies, as well as potential for orthographic interference. This could form part of an intervention study to investigate the effects of explicit instruction on promoting metacognitive strategies and awareness of orthographic influence on L2 pronunciation.

Relatedly, attention has been drawn to the need for more research with underrepresented language learners, such as adults with limited schooling in their first language. The motivation for increased attention to this population of language learners is ethical, theoretical and pedagogical. Language learners who are not represented in research are likely underserved and potentially harmed by inappropriate generalisations made about routes and rates of acquisition. Due to the reliance on L1 orthographic knowledge and difficulty mitigating cross-linguistic interference, reported for lower proficiency learners in this study, there is reason to be cautious around the quantity of written input during early language development. In the case of adults with limited L1 literacy, heavy reliance on L2 written input is likely to disadvantage those already most marginalised in the classroom. The importance placed on early phonological development by theoretical accounts also emphasises the need to understand orthographic influence during the earliest stages of L2 acquisition. Insight into early L2 phonological processes alongside emerging adult literacy is scarce, meaning that this route of enquiry would be valuable in relation to how we understand L2 phonological processes in general, as well as improving support for an underrepresented population of language learners.

As was highlighted in section 8.5, an extended version of the current study has already been built and piloted in Gorilla, including additional measures of individual differences, perception and production tasks, and additional stimuli. These elements address several limitations within the present study, including the fact it was designed to be a multi-session experiment. The circumstances which made this design difficult to implement have since changed, particularly Covid-19 measures which restricted in-person data collection. In-person or hybrid options would facilitate a longer experimental session and to include participants with more diverse literacy and educational backgrounds. The benefits of using an online platform, such as Gorilla, means the experiment can be run on multiple devices and taken to wherever is convenient for participants, including their language classrooms. One aspect that has not been piloted is the integration of eye-tracking. The presentation and testing phases could be easily adapted for a visual world paradigm, making this an exciting avenue to pursue. In particular, the inclusion of eye-tracking would shed light on orthographic influence on online learning processes.

Chapter 9: Conclusion

"All representations are wrong but some are useful" (Hoppe et al., 2022, p. 2243)[56]

Representations, by definition, are abstract and symbolic; offering a likeness of an external reality. The levels of representation discussed in this thesis include lexical, phonological and orthographic representations in the mental lexicon, as well as written representations of spoken language. These levels of representation encode certain information to support the perceiving, processing, and producing of language, as well as generalising to novel linguistic input. The human cognitive capacity is limited, both in terms of storage and processing, thus only some elements of the input to which we are exposed are taken in, often related to processes of competition, association, and prediction. For example, many phonetic and contextual details of a word may be omitted from a phonological representation, where frequency, saliency, redundancy, and reliability of cues in the input shape the information that is encoded (N. C. Ellis & Collins, 2009; Pierrehumbert, 2016). Each of these processes is tightly connected to previous experiences with language learning and individual cognitive capacity. Similarly, lexical representations, which at their core involve form-meaning mappings, are also developed from such input-related and individual factors. For those who are literate, these form-meaning mappings often include phonological and orthographic information. Orthographic or written representations then only partially denote the phonological information of a spoken language (Cutler, 2015; Ziegler & Goswami, 2005). Additionally, cross-modal information can be unreliable; such as silent letters, one-to-many mappings and sounds omitted from spelling. Thus, to some extent, all these representations are inaccurate, in comparison to the varied linguistic events which they represent.

That is not to say that all representations are minimal, as there is plenty of evidence that linguistic representations are highly detailed, with lexical quality and precision of representation being important contributors to language processing (Gor et al., 2021; Perfetti & Hart, 2002; Pierrehumbert, 2016; L. Verhoeven et al., 2019). In fact, as has been argued throughout this thesis, the lack of a sufficiently precise phonological representation of confusable L2 contrasts is central to the difficulty participants encounter when detecting

[56] The authors adapt this quote from George E. P. Box (Box, 1976)

correct form-meaning mappings in the present study. Furthermore, the perceptual salience of target contrasts and quality of phonological representations mediate the influence of orthographic input on the lexical encoding of novel words. The present study builds on a growing body of evidence that orthographic input may facilitate L2 lexical encoding of novel words where (1) phonological representations of the word are robust (Cutler, 2015; Escudero, 2015), (2) orthographic forms are sufficiently familiar to be decoded to corresponding units of phonology (Escudero et al., 2008; Hao & Yang, 2021; Shepperd, 2018), and (3) phoneme-grapheme correspondences (GPCs) across shared scripts are congruent (Hayes-Harb et al., 2010; Hayes-Harb & Cheng, 2016; Showalter, 2018). In these cases, orthographic representations in the lexicon can be considered potentially "useful" to L2 spoken word recognition. In contrast, exposure to written forms (1) mapping to imprecise phonological representations, (2) in an entirely unfamiliar script (Mathieu, 2016; Showalter & Hayes-Harb, 2015), or (3) which have incongruent GPCs across a shared script in the known and target languages (Cerni et al., 2019; Escudero et al., 2014; Rafat, 2016) are unlikely to result in "useful" orthographic representations. However, "usefulness" is a relative concept, and there is a distinction between a lack of usefulness (i.e., redundancy) and a biasing or interfering influence.

The present study exemplifies both these points. Firstly, exposure to any written input has an inhibitory rather than null effect on lexically encoding of confusable contrasts, in comparison to audio-only exposure. This is found both when written forms are in a shared script, where GPC mappings are incongruent (Arabic spelling), and in a distinct script, with reliable one-to-one GPCs (English spelling). The opposite is true for novel words differing by a well-established phonological contrast, when both shared and distinct written forms are reliable, leading to a facilitative effect. However, by using a mixed methods approach and including participants' perspectives, it is clear that written forms are broadly considered "useful", regardless of the perceptual salience of the contrast. This sense of "usefulness" is related to feeling confident using the novel forms, storing the novel forms in memory, forming associations, and improving the speed of learning. It is indeed the case that certain memory benefits are associated with learning from written over auditory input (Escudero et al., 2022; Nelson et al., 2005). However, the question of *quality* of representation remains, as the overshadowing of phonetic cues by orthographic input and the incongruence of cross-modal information can undermine the quality of lexical representations.

The strength of orthographic influence varies with proficiency, where those with lower proficiency in the target language rely more heavily on L1 orthographic knowledge and exhibit more bias from written input when learning novel forms. As proficiency improves, so does the ability to mitigate interference effects from orthographic input. This may be due to increasing stability and precision of lexical representations, in terms of both phonological and orthographic information, as well as increased cognitive control (Veivo et al., 2018; Veivo & Jarvikivi, 2013). Previous learning experiences and beliefs about the importance of written input are also important individual factors to acknowledge, alongside proficiency and cognitive ability. Learning experience includes becoming accustomed to specific script input as proficiency increases, where transliteration into the L1 script may be perceived as incongruous with L2 spoken forms (Hao & Yang, 2021), suggested by the preference for, and general advantage of, English spelling in the present study.

In sum, like the author quoted at the start of this thesis (Shafak, 2020), the learners in the present study are driven by the pursuit for meaning and reduction of uncertainty around form-meaning mappings encountered in the input. Sense-making and anchoring of the input they are exposed to takes place through idiosyncratic associations, visualisations, and the directing of attention. Personalised strategies are then employed to help overcome obstacles and support the learning of new forms. Participants' performance, beliefs, and strategies reflect the consistent use of orthographic knowledge and input to clarify and store new words, particularly where acoustic cues are less reliable. In contrast, Shafak's (2020) reflections arose from wrestling with the unreliability of orthographic cues. In her case, she did not face issues around perception and imprecise phonological representations of the words she was learning to write, as her anecdote related to childhood L1 acquisition. Meanwhile, the adult L2 learners presented here demonstrate the persistent and pervasive influence of orthographic influence on L2 phonology, reaching across writing systems and proficiency levels. Thus, this thesis joins with calls for caution and criticality when relying heavily on written input, particularly in the earlier stages of adult language learning.

Appendix I: Arabic information and consent

Information sheet

Learning words in a second language with and without written forms

Louise Shepperd is a researcher in learning English as a second language. Louise focuses specifically on studying how Arabic-speaking adults learn new English words and sounds. If Arabic is your first/main language, I would like to invite you to participate in this research project. You do not need to know any English in order to participate, all Arabic-speakers are welcome including beginners. ***Please read the following information carefully and let me know if anything is unclear or you would like more information before you decide to participate in this study. To learn more about general data protection laws you can ask the researcher.***

What is the purpose of the project?
The study investigates whether learning the spelling of new words influences the way that sounds are learned. It compares the influence of Arabic and English written forms to just hearing the words.

Do I have to take part?
Your participation is voluntary and you can stop at any time.

What do I have to do?
You will:

- *complete a short test for English (beginners can skip)*
- *complete a background information and language questionnaire*
- *learn 12 invented words based on English pronunciation*
- *play a games to see whether you can remember the 12 words*
- *complete a questionnaire about your experience of learning the words*

This should take around 30 mins to complete. You can use the link in your email to re-enter if you need to take a break or come back later.

After completing the study, you will be given access to English pronunciation videos created by the lead researcher. You can also choose to add your name to a leaderboard with other participants. The top score will be whoever learns the words with the highest accuracy and in the fastest time. The top three scores will win 2 hours of one-to-one language classes with the lead researcher.

Will I be recorded, and how will my work be used?
There will be no voice or video recordings of your activities made during this research. Only your questionnaire answers and responses during word learning and testing will be used for analysis and future publications.

What if something goes wrong?

If you feel something has gone wrong or would like to raise an issue/complaint, you are advised to contact LOUISE SHEPPERD by email louise.shepperd@york.ac.uk, or the Chair of Ethics Committee via email education-research-administrator@york.ac.uk. If your complaint relates to how your personal data has been handled, you can contact the University's Data Protection Officer at dataprotection@york.ac.uk

Will my taking part in this project be kept confidential?

Your participation in the project will be kept strictly confidential. The data that you provide (e.g. questionnaire responses, test results etc.) will be stored by a code and two months after completion will be fully anonymised. So, you are free to withdraw your data during the first two months after data collection, but after then it will not be identifiable.

Data will be stored on a password protected computer. I support open access to scientific data which means anonymised data will be managed professionally and stored indefinitely with the University's Research Data York service.

What will happen to the results/findings?

Anonymised results and findings will form part of a PhD thesis at the University of York and are likely to be shared at academic conferences, published in scientific journals and stored in University archives.

Who is organising and funding the research?

This research is funded by ESRC White Rose Doctoral Training Partnership.

Who has ethically reviewed the project?

This project has been ethically approved via the Department of Education's ethics review procedure.

Contact for further information

If you have any questions about the study, please contact:

Email: Louise Shepperd

Thank you for taking the time to read about the project!

Taking part in the project

I have understood the information given to me about this research project.

- ○ *Yes*
- ○ *No*

I understand that my taking part is voluntary and that I can withdraw from the study at any point.

○ *Yes*

○ *No*

I have normal or corrected vision and hearing (e.g. glasses, contact lenses, hearing aid).

○ *Yes*

○ *No*

I have known cognitive impairments or diagnosed dyslexia.

○ *Yes*

○ *No*

How my information will be used during and after the project

I understand that my data will not be identifiable and the data will be used for a PhD project at the University of York, as well as subsequent publications, presentations.

○ *Yes*

○ *No*

I give permission for the responses I provide on all tasks/activities to be stored securely on the University of York's Google Drive

○ *Yes*

○ *No*

☐ *I consent to take part in this study*

ورقة المعلومات

تعلم الكلمة بلغة ثانية مع أو بدون أشكال مكتوبة

لويز شبرد هي باحثة في تعلّم اللغة الإنجليزية كلغة ثانية. تركّز لويز تحديدًا على دراسة كيفية تعلّم البالغين من الناطقين باللغة العربية كلمات وأصوات باللغة الإنجليزية الجديدة. إذا كانت اللغة العربية هي لغتك الأساسية، أود أن أدعوك للمشاركة في هذا المشروع البحثي. لا تحتاج إلى معرفة أي كلمة انجليزية من أجل المشاركة نرحب بجميع متعلمي اللغة الانجليزية من المتحدثين العرب، بما فيهم المبتدئين. **الرجاء قراءة المعلومات التالية بشكل جيد وأخبرنا إذا كان هناك أي شيء غير واضح و تريد معلومات أكثر قبل أن تقرر المشاركة في البحث. لتعرف أكثر عن لائحة حماية البيانات العامة، يمكنك أن تسأل الباحثة.**

ما هو الهدف من المشروع؟

تبحث الدراسة فيما إذا كان رؤية الكلمات المكتوبة يؤثر على طريقة تعلم الأصوات الجديدة. يقارن تأثير النماذج المكتوبة باللغتين العربية والإنجليزية بسماع الكلمات فقط.

هل يجب على المشاركة؟

مشاركتك تطوعية بالكامل ويمكنك تغيير رأيك والتوقف في أي وقت.

ماذا علي أن أفعل؟

سوف تقوم بعمل التالي

- اختبار قصير للغة الإنجليزية (يمكن للمبتدئين تخطيه).
- استبيان حول المعلومات الخلفية واللغة.
- تعلّم 12 كلمة مخترعة تعتمد على نطق اللغة الإنجليزية.
- لعبة لاختبار مدى تذكرك للكلمات الاثنتي عشرة.
- استبيان حول تجربتك في تعلم الكلمات.

سوف يستغرق ذلك حوالي 30 دقيقة . يمكنك استخدام الرابط في بريدك الإلكتروني لإعادة الدخول إذا كنت بحاجة إلى أخذ استراحة أو العودة في وقت لاحق.

بعد الانتهاء من الدراسة، سيتم تزويدك من الباحث مقاطع فيديو عن النطق في اللغة الانجليزية. يمكنك أيضا اختيار إضافة اسمك إلى اللوحة الرئيسية مع المشاركين الآخرين. الدرجة الأعلى ستكون لمن يتعلم الكلمات بأعلى دقة وبأسرع وقت. ستفوز الدرجات الثلاث الأولى بساعتين من دروس اللغة الفردية مع الباحث الرئيسي.

هل ستسجَّل إجاباتي وكيف سيتم استخدامها؟

لن يتم حفظ أي تسجيل صوتي أو مرئي في حال مشاركتكم في هذا البحث. ولن يتم استخدام أي تسجيل صوتي أو مرئي في تحليل النتائج أو في الدراسات المستقبلية. ما سيتم استخدامه وتحليله ونشره في البحوث العلمية هو إجاباتك على الاستبيانات واختبار الكلمات، علمًا بأنه سيتم الحفاظ على سريّة هويتك بعناية تامة لحماية خصوصيتك.

ماذا لو حدث خطأ ما؟

إذا كنت تشعر بأن شيئًا ما قد حدث خطأ أو ترغب في تقديم شكوى، يُنصح بالتواصل عبر البريد الإلكتروني مع لويز شبرد على louise.shepperd@york.ac.uk.

أو رئيس لجنة الأخلاقيات عبر البريد الإلكتروني على education-research-administrator@york.ac.uk

إذا كانت شكواك تتعلق بكيفية التعامل مع بياناتك الشخصية ، فيمكنك الاتصال بمسؤول حماية البيانات بالجامعة على dataprotection@york.ac.uk

هل ستبقى مشاركتي في هذا المشروع سرية؟

ستبقى مشاركتك في المشروع سرية للغاية. سيتم تخزين البيانات التي تقدمها (مثل الردود على الاستبيان ونتائج الاختبار وما إلى ذلك) بواسطة شفرة وبعد شهرين سيتم إخفاء هويتها بالكامل. لذلك، لك مطلق الحرية في سحب بياناتك خلال الشهرين الأولين بعد جمع البيانات، ولكن بعد ذلك لن يتم التعرف عليها.

نؤكد لك بأنه سيتم تخزين البيانات في خزائن آمنة وعلى جهاز كمبيوتر محمي بكلمة مرور. أنا أؤيد الوصول الحر للبيانات العلمية مما يعني أنه سيتم إدارة البيانات مجهولة المصدر بشكل احترافي وتخزينها إلى أجل غير مسمى مع خدمة الجامعة لبيانات الأبحاث.

ماذا سيحدث للنتائج؟

ستشكل النتائج مجهولة المصدر جزءًا من أطروحة دكتوراه في جامعة يورك ومن المرجح أن يتم مشاركتها في المؤتمرات الأكاديمية ونشرها في المجلات العلمية وتخزينها في أرشيف الجامعة

من يقوم بتنظيم وتمويل البحث؟

تم تمويل هذا البحث من قبل ESRC White Rose Doctoral Training Partnership

من قام بمراجعة المشروع أخلاقيا؟

تمت الموافقة على هذا المشروع أخلاقيا من قسم مراجعة الأخلاقيات في كليه التعليم

معلومات التواصل

إذا كان لديك أي أسئلة حول الدراسة ، فيرجى الاتصال بلويز شيبرد عبر البريد الإلكتروني: louise.shepperd@york.ac.uk
نشكرك على الوقت الذي قضيته في القراءة عن المشروع

يرجى وضع علامة في كل مربع إذا كنت سعيدًا بالمشاركة في هذا البحث

أؤكد أنني قد قرأت وفهمت المعلومات التي قُدمت لي حول المشروع البحثي أعلاه

- نعم
- لا

أفهم أن المشاركة في هذه الدراسة تطوعية

- نعم
- لا

اسمع وأرى بشكل جيد أو مُصحح (مثل استعمال النظارات الطبية-العدسات اللاصقة-سماعات مُساعدة)

- نعم
- لا

ليس لدي أي إعاقات إدراكية معروفة أو تشخيص عسر الحركة.

- نعم
- لا

كيف سيتم استخدام معلوماتي أثناء المشروع وبعده

أعرف أن بياناتي لن تكون معرفة ويمكن استخدام البيانات في المنشورات والعروض التقديمية

- نعم
- لا

أعطي الإذن لاستخدام الإجابات التي أقدمها بجميع المهام/الأنشطة أن يتم تخزينها بشكل آمن على محرك جوجل التابع لجامعة يورك

- نعم
- لا

☐ **أنا أوافق على المشاركة في هذه الدراسة**

Appendix II: English information and consent

Participant Information Sheet

Louise Shepperd is researching second language learning, specifically how new words and sounds in English are learned by Arabic-speaking adults compared to English-speaking adults. If English is your first or main language and you do not know any Arabic (or languages that use the Arabic writing system), I would like to invite you to take part in this research project. **Please read this information carefully and let us know if anything is unclear or you would like more information. For information about General Data Protection Regulation (GDPR) please ask the researcher for more information.**

1. What is the aim of the project?

The study investigates whether learning the spelling of new words influences the way that sounds are learned. It compares the influence of Arabic and English written forms to just hearing the words.

2. Do I have to take part?

Your participation is voluntary and you can stop at any time.

3. What do I have to do?

You will:

- complete a background questionnaire
- learn 12 invented words based on English pronunciation
- try to remember the 12 words in a short test
- complete a questionnaire about your experience of learning the words

This should take around 30 mins to complete. You can use the link in your email to re-enter if you need to take a break or come back later.

After completing the study, you can choose to enter your name on the leaderboard. You will then be emailed to see if your score on the short test was in the top twenty.

4. Will I be recorded, and how will my work be used?

There will be no voice or video recordings of your activities made during this research. Only your questionnaire answers and responses during word learning and testing will be used for analysis and future publications.

5. What if something goes wrong?

If you feel something has gone wrong or would like to raise an issue/complaint, you are advised to contact LOUISE SHEPPERD by email louise.shepperd@york.ac.uk, or the Chair of Ethics Committee via email education-research-administrator@york.ac.uk. If your complaint relates to how your personal data has been handled, you can contact the University's Data Protection Officer at dataprotection@york.ac.uk

6. Will my taking part in this project be kept confidential?
Your participation in the project will be kept strictly confidential. The data that you provide (e.g. questionnaire responses, test results etc.) will be stored by a code and two months after completion will be fully anonymised. So, you are free to withdraw your data during the first two months after data collection, but after then it will not be identifiable.

Data will be stored on a password protected computer. I support open access to scientific data which means anonymised data will be managed professionally and stored indefinitely with the University's Research Data York service.

7. What will happen to the results/findings?
Anonymised results and findings will form part of a PhD thesis at the University of York and are likely to be shared at academic conferences, published in scientific journals and stored in University archives.

8. Who is organising and funding the research?
This research is funded by ESRC White Rose Doctoral Training Partnership.

9. Who has ethically reviewed the project?
This project has been ethically approved via the Department of Education's ethics review procedure.

10. Contact for further information
If you have any questions about the study, please contact:
Email: Louise Shepperd
Thank you for taking the time to read about the project!

Taking part in the project

I have understood the information given to me about this research project.

- ○ Yes
- ○ No

I understand that my taking part is voluntary and that I can withdraw from the study at any point.

- ○ Yes
- ○ No

I have normal or corrected vision and hearing (e.g. glasses, contact lenses, hearing aid).

- Yes
- No

I have known cognitive impairments or diagnosed dyslexia.

- Yes
- No

How my information will be used during and after the project

I understand that my data will not be identifiable and the data will be used for a PhD project at the University of York, as well as subsequent publications, presentations.

- Yes
- No

I give permission for the responses I provide on all tasks/activities to be stored securely on the University of York's Google Drive

- Yes
- No

- [ ] I consent to take part in this study

Appendix III: English and Arabic debriefing information

Thank you again for your participation in my doctoral research. In case it is of interest, here is some more information about the study you have contributed to. This study investigates how Arabic and English written forms influence Arabic-speakers when they learn words containing difficult English sounds, like /f/ and /v/. In addition, the study explores what participants perceive to be most helpful and how that compares to their scores on the word learning test.

Previous research
Previous studies have found both positive and negative effects of written forms on learning new sounds and words, and often warn against overreliance on written forms when learning new words. To date, only a few studies have investigated the influence across different writing systems and rarely include participant reflections.

Present study
All participants are asked to memorise the same 12 invented words and images, which began with the sounds /m/, /n/, /f/ or /v/. The sounds /m/ and /n/ exist in English and Arabic, and are spelled with two different letters in each script. In contrast, the sound /v/ is not commonly found in Arabic and can be difficult to distinguish from /f/. This is reflected in Arabic spelling, which typically spells /v/ with the same letter as /f/ (i.e. <ف>), whereas English spells the two sounds with two letters (i.e. <f> and <v>).

After memorising the words, accompanied by English spelling, Arabic spelling or no spelling, participants have to distinguish between words which only differ by the first sound (e.g. <famel> vs. <vamel>). It is normal to find this test difficult and is not a reflection of your memory or ability to learn languages. The speed and accuracy of responses will be analysed to see whether scores differ in relation to different spellings, questionnaire responses about learning strategies, literacy usage and English proficiency. The responses of those who complete the same test again after 24 hours will then offer more insight into learning processes after a sleep cycle.

Findings
These findings will shed light on the impact of written language when learning new sounds and words, including transcription in the first language writing system. It also hopes to reveal assumptions around literacy in word learning and pronunciation.

I will contact you after the study is completed with information about the final results. In the meantime, feel free to contact me with any questions or feedback about the study. I can also provide more information about use and storage of the data you have contributed towards this research (louise.shepperd@york.ac.uk).

Thank you very much for your time and interest.

Further reading

Hayes-Harb, R. & Barrios, S. (2021). The influence of orthography in second language phonological acquisition. Language Teaching 54, 297-326.

أجدد لك الشكر على مساهمتك في دراستي البحثية لرسالة الدكتوراه. إن كانت لديك الرغبة في معرفة المزيد عن البحث، مرفق لك بعض المعلومات عن هذه الدراسة التي شاركت فيها.

أبحث في هذه الدراسة عن تأثير قراءة الكلمة مكتوبة عند تعلّم ناطقين اللغة العربية كلمات جديدة تحتوي على أصوات غير مألوفة باللغة الإنجليزية، مثل التفريق بين/f/ و /v/. كما تهتم هذه الدراسة في رأي المشاركين عن مدى الاستفادة من وجود الكلمة مكتوبة بالمقارنة مع عدم وجودها، وتأثير ذلك على نتائجهم في اختبار تعلّم الكلمات.

الدراسات السابقة

تعارضت نتائج الدراسات السابقة لتعلّم الكلمات ذات الأصوات غير المألوفة مع قراءة الكلمة مكتوبة بين وجود أثرًا إيجابي في بعض الدراسات وأثرًا سلبي في البعض الآخر. كما حذّرت بعض الدراسات من فرط الاعتماد على تعلّم الكلمات بقراءة الكلمة مكتوبة. حتى الآن، لا يوجد غير دراسات قليلة بحثت في هذا الموضوع باستخدام أنظمة كتابية مختلفة، مثلًا عربية ولاتينية، والأندر من ذلك أن تشمل الدراسة رأي المشاركين حول الفرق بين تعلّم الكلمة مكتوبة أم مسموعة.

هذه الدراسة

يتطلب من المشاركين في هذه الدراسة حفظ ١٢ كلمة مخترعة مع صورة لكل كلمة، هذه الكلمات تبدأ بأحد هذه الأصوات /m/ أو /n/ أو /f/ أو /v/. الأصوات التالية موجودة باللغتين العربية والانجليزية: /m/ و /n/ و /f/، ولكن الصوت /v/ غير موجود باللغة العربية بالتالي يصعب التفريق بينه وبين الصوت /f/. ينعكس ذلك على تهجئة الكلمات باللغة العربية، الذي ينتج عن تهجئة /v/ بنفس صوت /f/ (أي <ف> لكلا الصوتين)، بينما ينتج عن تهجئة الصوتين باللغة الإنجليزية بحرفين مختلفين كتابيًا (أي <f> و <v>).

بعد حفظ الكلمات بمختلف الطرق: مع كتابة إنجليزية، أو مع كتابة عربية، أو بدون كتابة، يتطلب أيضًا من المشاركين في هذه الدراسة التمييز بين الكلمات التي تختلف في الصوت الأول فقط (مثلًا <famel> مقابل <vamel>). من الطبيعي أن تواجه صعوبة في هذا الاختبار، ولا ينعكس ذلك على مدى قوة ذاكرتك أو قدرتك على تعلّم اللغات. سيتم تحليل سرعة الوقت المستغرق في الإجابة بالإضافة إلى دقتها لإيجاد الفروقات في النتائج بين طرق التعلّم المختلفة، وبين ردود وآراء المشاركين في الاستبيان حول طرق التعلّم، وبين درجة المشاركين العلمية ومدى اتقانهم للّغة الإنجليزية كلغة ثانية. عند اكمال المشاركين نفس الاختبار في اليوم التالي، سيتسنى لي معرفة المزيد عن أثر عمليات التعلم بعد دورة نوم كاملة.

النتائج

هذه النتائج ستسلّط الضوء على أثر اللغة الكتابية عند تعلّم كلمات جديدة وأصوات غير مألوفة، بالإضافة إلى تعلّم كلمات جديدة بلغة ثانية باستخدام نظام الكتابة في اللغة الأم. كما نأمل أن يساهم هذا البحث للتوصّل إلى فهم للافتراضات حول معرفة القراءة والكتابة في تعلّم ونطق الكلمات.

سأتواصل معك عند اكتمال هذه الدراسة لأطلعك على نتائجها. في الوقت الحالِ، أرجو عدم التردد في التواصل معي إذا كان لديك أي استفسار أو اردت التعليق على هذه الدراسة. يمكنني أيضًا توفير المزيد من المعلومات عن طرق استخدام وتخزين بياناتك التي ساهمت بها في هذه الدراسة (louise.shepperd@york.ac.uk).

اشكر لك وقتك ومساهمتك.

لقراءة المزيد

Hayes-Harb, R. & Barrios, S. (2021). The influence of orthography in second language phonological acquisition. Language Teaching 54, 297-326.

Appendix IV: English and Arabic plain language summaries

The influence of written forms when learning words with difficult sounds: Arabic-speaking learners of English as a second language[57]

What this research was about and why it is important

Many adult language learners are exposed to large quantities of written language from the earliest stages of learning a new language, particularly in classroom settings. This is in stark contrast to child language acquisition, where exposure to, and proficiency with, spoken language precedes literacy development. Adult language learners regularly face obstacles around perceiving and producing the sounds of a new language, which are not found in their first or known languages. The present study sought to investigate whether written forms help or hinder learning new words that differ by confusable sounds. Also does the influence differ if words are taught with spellings from different writing systems?

What the researchers did

- 114 Arabic-speakers (varying English proficiency) and 117 English-speakers (no Arabic experience) completed an online word learning and matching task.
- Participants learned 12 artificial words, accompanied by a novel image and (1) Arabic spelling, (2) English spelling, or (3) audio only.
- Words differed by an easy (i.e., /m-n/) or difficult (i.e., /f-v/) sound pair.
- Letters mapped to sounds for both pairs in English: <m-n> and <f-v>.
- Letters mapped to sounds for easy but not difficult sounds in Arabic: <م- ن> and <ف>.
- Participants were then asked to reflect on their learning with and without different written forms and what strategies they used to learn the words

What the researchers found

- Arabic-speaker accuracy in the matching task was worse for difficult /f-v/ words with both English and Arabic spelling, compared to audio only.
- The opposite was true for easy /m-n/ words.
- The vast majority of participants believed that seeing the written forms helped them, particularly English spelling.
- Strategies using the sounds and written forms of a language did not improve accuracy, but evaluative strategies did offer some benefits.
- Arabic-speaking participants with higher English proficiency relied less on Arabic literacy and were better able to avoid first language interference.

Things to consider

[57] Developed following the OASIS summaries template (Marsden et al., 2018).

- There may be advantages associated with written forms, such as rapid memorisation of vocabulary. However, reliance on written forms interferes with encoding sufficient detail about sounds in memory, making it difficult to recognise spoken forms. This is particularly important when learning words containing confusable sounds.

تأثير الصيغة الكتابية على تعلم الكلمات ذات الأصوات الصعبة: متعلمي اللغة الأنجليزية كلغة ثانية من متحدثي اللغة العربية.

ماهو هذا البحث وماهي أهميته:

العديد من متعلمي اللغة الأنجليزية البالغين يتعرضون لكمية كبيرة من الصيغ المكتوبة من المراحل المبكرة من تعلم اللغة الجديدة خصوصاً في الفصول الدراسية وهذا يختلف بشكل كلي مع طريقة اكتساب اللغة عند الأطفال من ناحية التعرض للغة والطلاقة فيها حيث أن الأطفال يتعلمون اللغة المنطوقة قبل التمكن من القراءة والكتابة في نفس اللغة.
المتعلمين البالغين عادة ما يواجهون في فهم ونطق الأصوات الجديدة في اللغة المراد تعلمها والتي تكون غير موجودة في لغتهم الأم.
الدراسة الحالية تهدف إلى التحقق مما إذا كانت الصيغة الكتابية قد تساعد او تعيق تعلم الكلمات الجديدة والتي تختلف لاحتوائها على أصوات مربكة للمتعلمين. وأيضا هل هناك اختلاف في تأثير تدريس الكلمات مع صيغتها الإملائية والتي تكون من نظام كتابي مختلف عن اللغة الأم.

ماذا عمل الباحثين:

- ١١٤ متحدث للعربية (مع اتقان متفاوت للغة الإنجليزية) و ١١٧ متحدث بالإنجليزية (بدون معرفة مسبقة باللغة العربية) أكملوا مهمة تعلم كلمات وتوصيلها اونلاين.
- المشاركين تعلموا ١٢ كلمة غير حقيقية و صورة غير مألوفة مع إما صيغة إملاء واحدة باللغة العربية أو صيغتي املاء باللغة الإنجليزية أو صوت فقط.
- تختلف الكلمات باختلاف زوج من الأصوات السهلة مثل (m-n) أو زوج من الأصوات الصعبة مثل (f-v).
- في اللغة الإنجليزية الحروف المستخدمة لتمثيل الأصوات كانت <m-n> و<f-v> في اللغة الإنجليزية.
- في اللغة العربية الحروف المستخدمة في الإملاء كان <م -ن> لكل من (m,n) ولكن تم استخدام <ف> فقط لتمثيل الصوتين (f,v) لعدم وجود (v) في اللغة العربية.
- طُلب من المشاركين التفكر في طريقة تعلمهم مع أو بدون الصيغ الإملائية و ما هي الاستراتيجيات التي استخدموها لتعلم الكلمات.

ماذا وجد الباحثين:

- في مهمة توصيل الكلمات، كان أداء متحدثي اللغة العربية أسوأ مع الأصوات الصعبة (f,v) في كلٍ من صيغة الإملاء العربي والانجليزي مقارنة بالصوت فقط.
- العكس كان صحيحاً للأصوات السهلة /m-n/.
- أغلبية المشاركين يعتقدون أن رؤية إملاء الكلمات ساعدهم خصوصاً الإملاء باللغة الإنجليزية.
- الاستراتيجيات تشمل معرفة الأصوات والصيغة المكتوبة للغة لم يساعد على تطويرها لكن تقييم الاستراتيجيات كان له بعض الفائدة.
- المشاركين من متحدثي اللغة العربية والذين لديهم طلاقة أكثر باللغة الإنجليزية كانوا أقل اعتماداً على الصيغة المكتوبة باللغة العربية و يظهر أنهم أفضل في تجاوز التداخل ما بين اللغتين.

نقاط مهمة للأخذ في الاعتبار:

ربما يوجد هناك أفضلية في استخدام الصيغة المكتوبة مثل الحفظ السريع للكلمات، ولكن الاعتماد الكبير على الصيغة المكتوبة قد يقيّد ترميز الأصوات في الذاكرة مما يسبب صعوبة في إدراك الكلام المنطوق. وهذا قد يكون مهماً في تعلّم الكلمات التي تحوي أصوات ذات صعوبة للمتعلمين.

Appendix V: Data pre-processing Rmd output

This document outlines the steps to process the accuracy and reaction time data from the word learning and audio-visual matching task, including exclusions. The structure of the document is as follows:

1. Description of demographic data
2. Description of task set up data
3. Description of task completion data
4. Calculation of exclusions and processed data

Demographic data

To begin with, the demographic data for Arabic and English-speaking participants was explored. An overview of the data is outlined by:

1. demographic information (age, gender, education etc.)
2. language background (L1, dialect, other languages etc.)
3. English language proficiency and experience (Arabic-speakers only)

The table below gives an overview of the mean age in years for the two participant groups, as well as gender and completed level of education.

Table V.1: Demographic overview for participants by L1

	L1 Arabic			**L1 English**		
	n	**Mean**	**SD**	***n***	**Mean**	**SD**
age	128	30.891	7.94	133	37.481	14.902
gender	128			133		
… female	94	73.4%		88	66.2%	
… male	33	25.8%		41	30.8%	
… not listed	1	0.8%		4	3%	
education	128			133		
… primary	1	0.8%		0	0%	
… secondary	16	12.5%		13	9.8%	
… other qual	2	1.6%		18	13.5%	
… bachelors	41	32%		40	30.1%	
… masters	46	35.9%		46	34.6%	
… doctorate	22	17.2%		16	12%	

Age is in years, gender 'n' is 'not listed', and education is completed qualifications

The figures below plot the spread and distribution of participants by demographic factors and language group. English participants were spread across a wider age range than Arabic participants, and it was also clear that Arabic-speakers were confused about how to interpret the question about dyslexia and cognitive impairments. The groups were broadly comparable in terms of gender, level of education and lack of vision or hearing impairments.

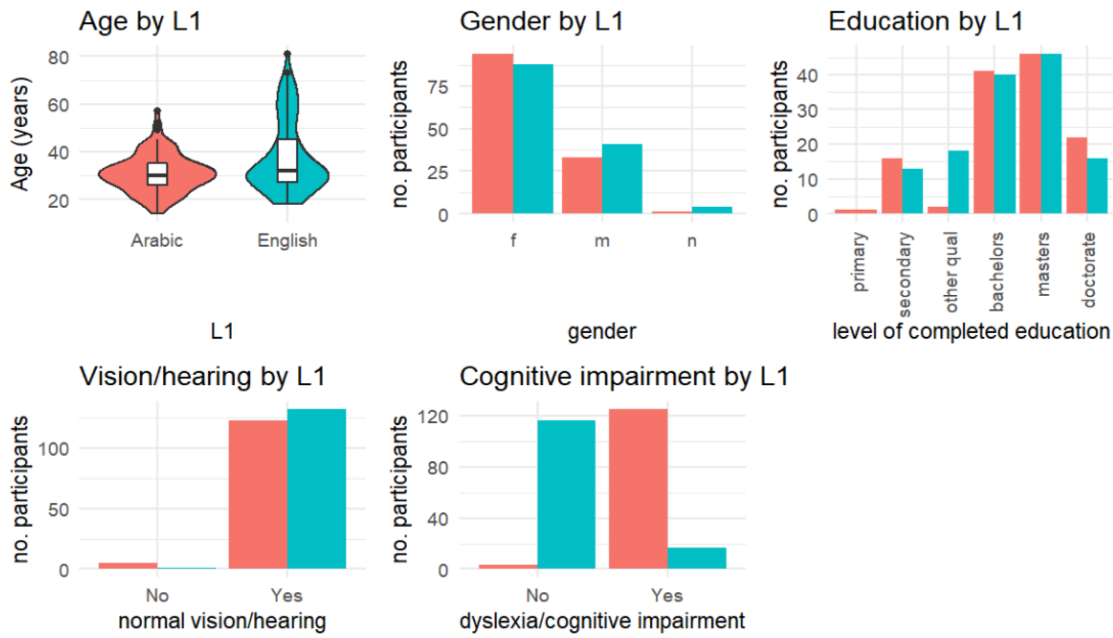


Figure V.1: Overview of participant demographic information for both language groups

The map below shows participant language and national background. Dialects mapped almost exactly onto reported nationalities, so we can see where most Arabic and English-speaking participants are from around the world, and what varieties they are likely to speak.

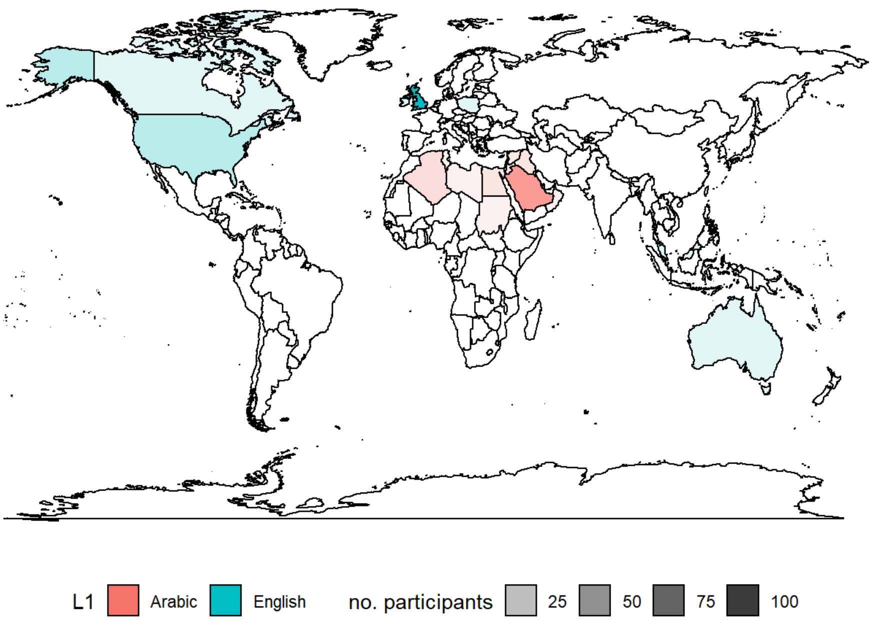


Figure V.2: Choropleth map of participant nationalities by L1 language background

The bar charts below provide further insight into the different dialects spoken by the L1 Arabic and English-speakers. To explore the language background of participants, Arabic dialects were grouped into five regional varieties.

1. **Maghrebi** e.g. Moroccan, Algerian, Tunisian, Libyan
2. **Egypto-Sudanic** e.g. Egyptian, Sudanese, Sa'adi, Chadian
3. **Mesopotamian** e.g. Iraqi dialects
4. **Levantine** e.g. Palestinian, Jordanian, Syrian, Lebanese
5. **Gulf/ Peninsula** e.g. Najdi, Gulf, Bahrani, Hejazi, Yemeni, Omani

The first figure demonstrates the bias towards speakers of Gulf dialects, most of whom were speakers of Saudi Arabian dialects. However, there was a spread across the principal Arabic varieties. An overview of English dialects shows most participants were speakers of British English varieties.

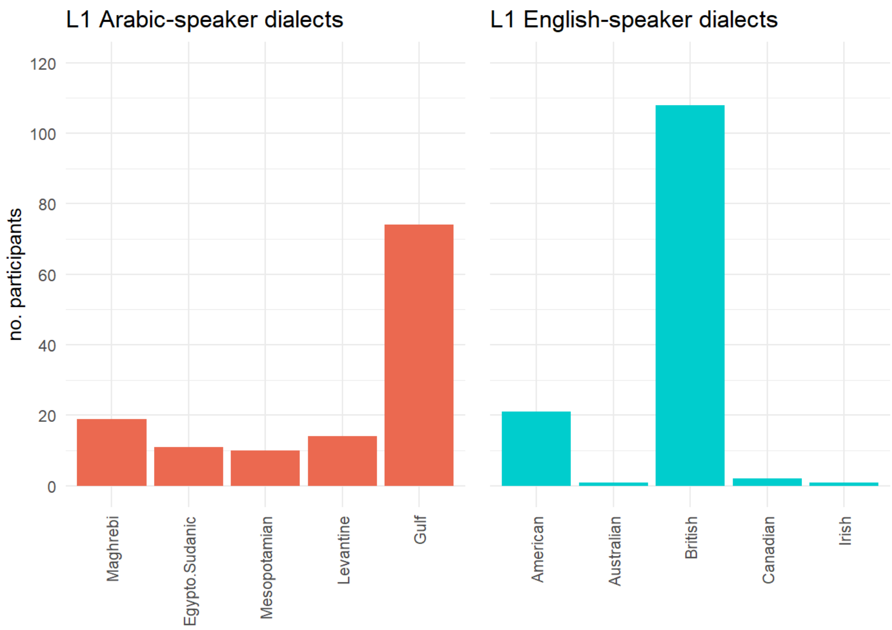


Figure V.3: Number of participants by dialect

To understand the extent to which participants were exposed to English and Arabic, they were asked to estimate how many hours they spent listening, reading and interacting in each language every day. Proportion of exposure was calculated as the total hours exposed to English or Arabic divided by the total number of hours exposed to all languages. English-speakers reported hours exposed to Arabic, English and other languages. Arabic speakers reported hours exposed to their colloquial dialect and MSA separately, to capture diglossic language usage.

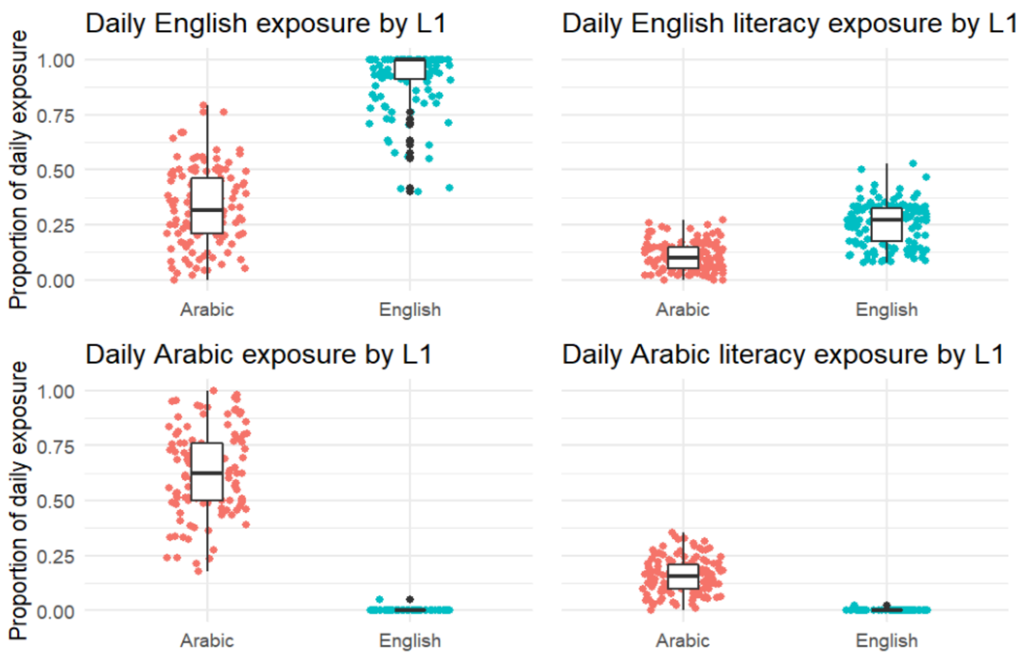


Figure V.4: Estimated daily exposure to English and Arabic language and literacy

The figures below detail the additional languages reported by participants. A large proportion of participants in both language groups reported not speaking additional languages. None of the English-speakers reported knowing any languages that were written with the Arabic alphabet. Meanwhile, a number of Arabic-speakers reported knowing French, Spanish, Turkish and Swedish, all of which are written with the Roman alphabet. The popularity of French is likely linked to the colonial influence of French in North Africa, and its influence on Maghrebi dialects.

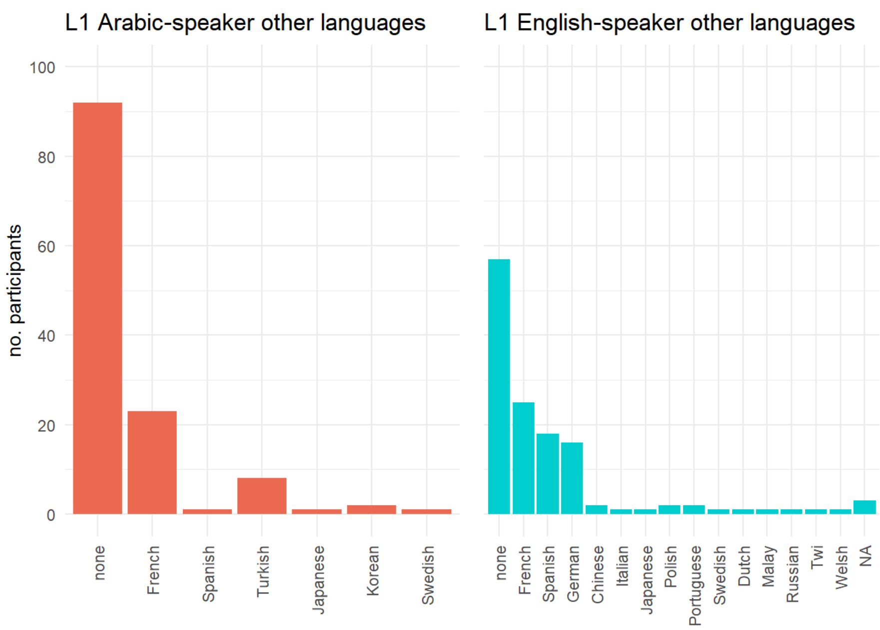


Figure V.5: Additional languages reported by participants from both language groups

Next, the following figures provide an overview of participants' ability and experience in English. Therefore, the data presented focuses on the Arabic-speaking participants.

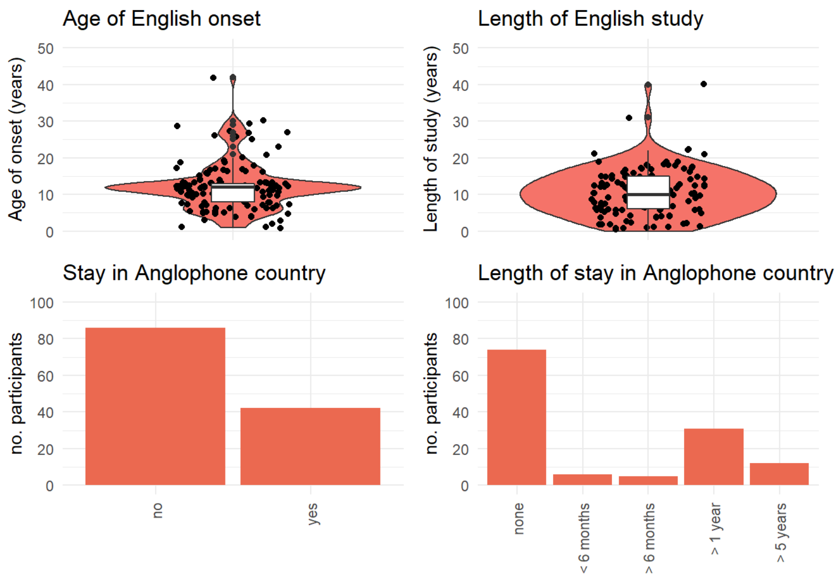


Figure V.6: Overview of L1 Arabic speaking participants and their experience with English

The figure below provides an overview of self-reported English level by Arabic-speaking participants, and demonstrates the large proportion of advanced English speakers.

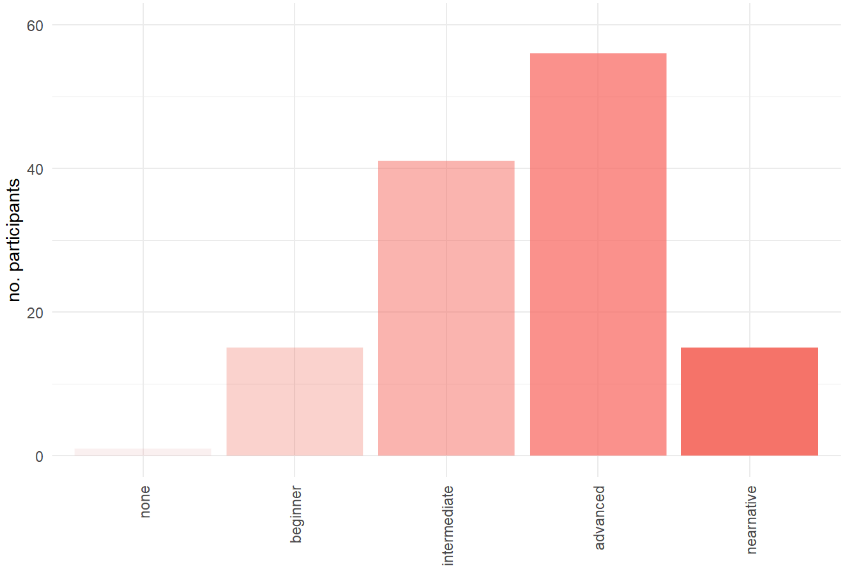


Figure V.7: L1 Arabic participants' self-reported proficiency level in L2 English

Further insight was offered by asking participants to self-report their ability across reading, writing, speaking and listening in English.

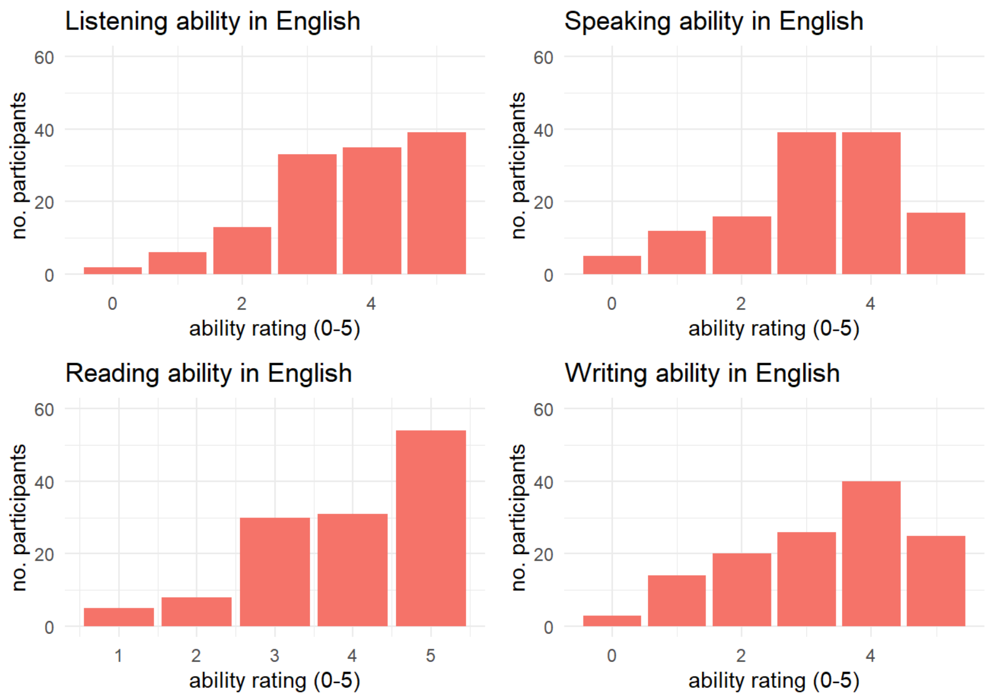


Figure V.8: L1 Arabic self-reported ability across the different skills of listening, speaking, reading and writing in L2 English

In order to triangulate self-reported proficiency data, participants also completed a short English proficiency test and their scores are plotted below.

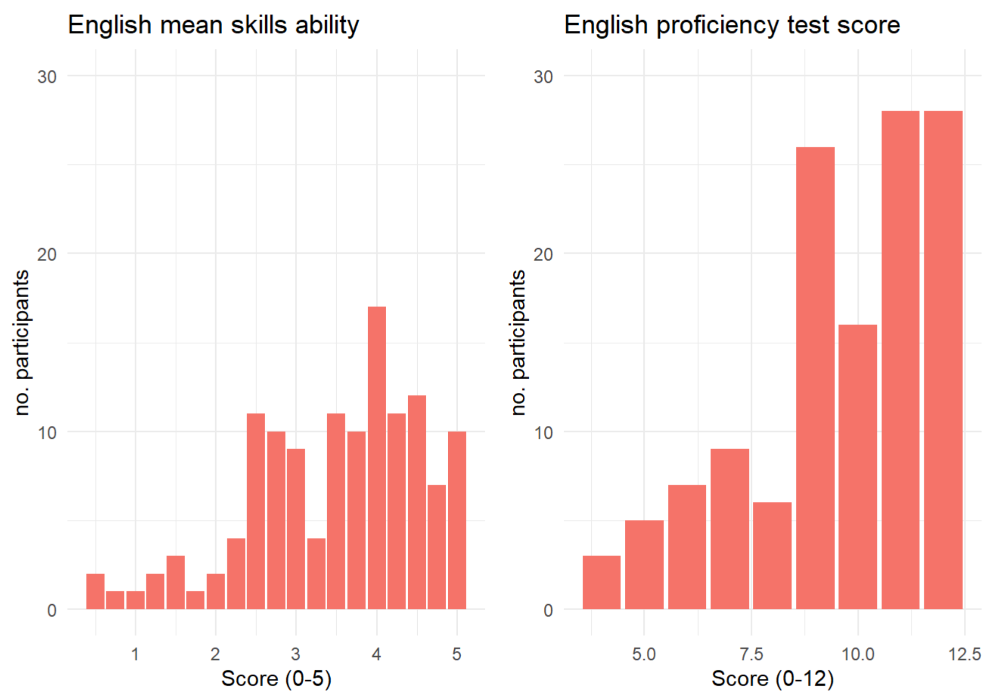


Figure V.9: L1 Arabic-speaker mean self-reported L2 English skills score and English proficiency test score

These indicators of proficiency were then explored for correlations with each other. The correlation matrix below includes the significant correlation coefficients for the variables that indicate participant English proficiency.

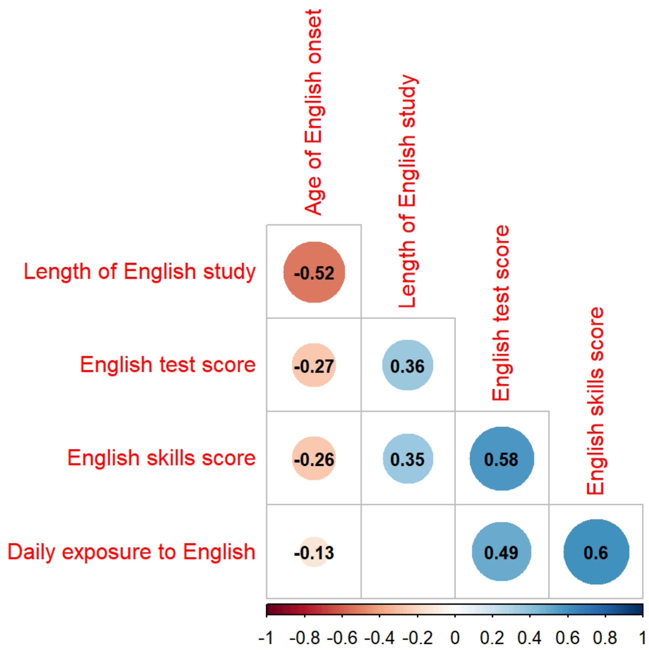


Figure V.10: Correlation matrix for English proficiency measures

In order to ensure that all English-speakers had little to no Arabic ability, the data was explored by reported proficiency. First of all, it was clear from the figure below that most English-speaking participants had not spent time living in an Arabic-speaking country. In addition, the vast majority reported no proficiency in Arabic.

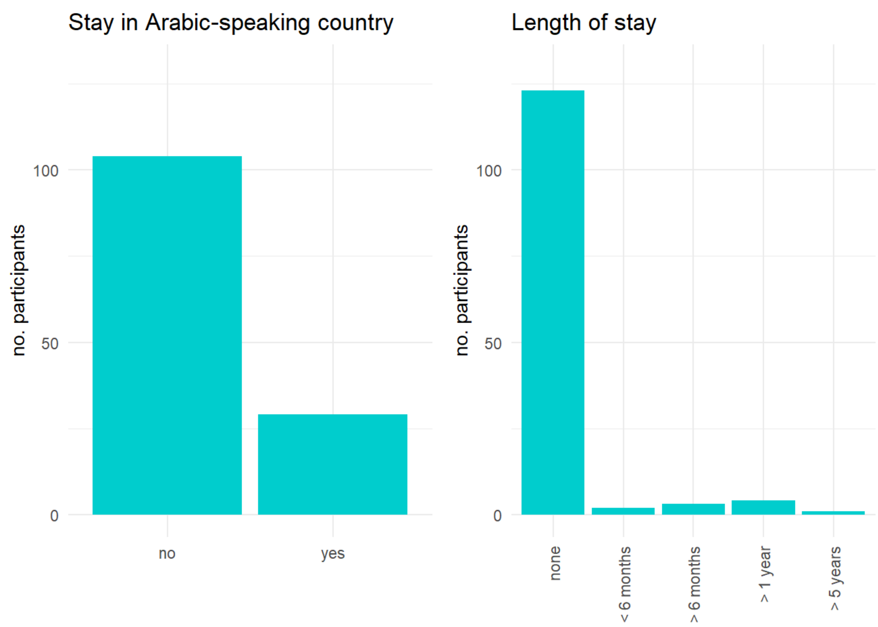


Figure V.11: L1 English participants' time spent in Arabic-speaking countries

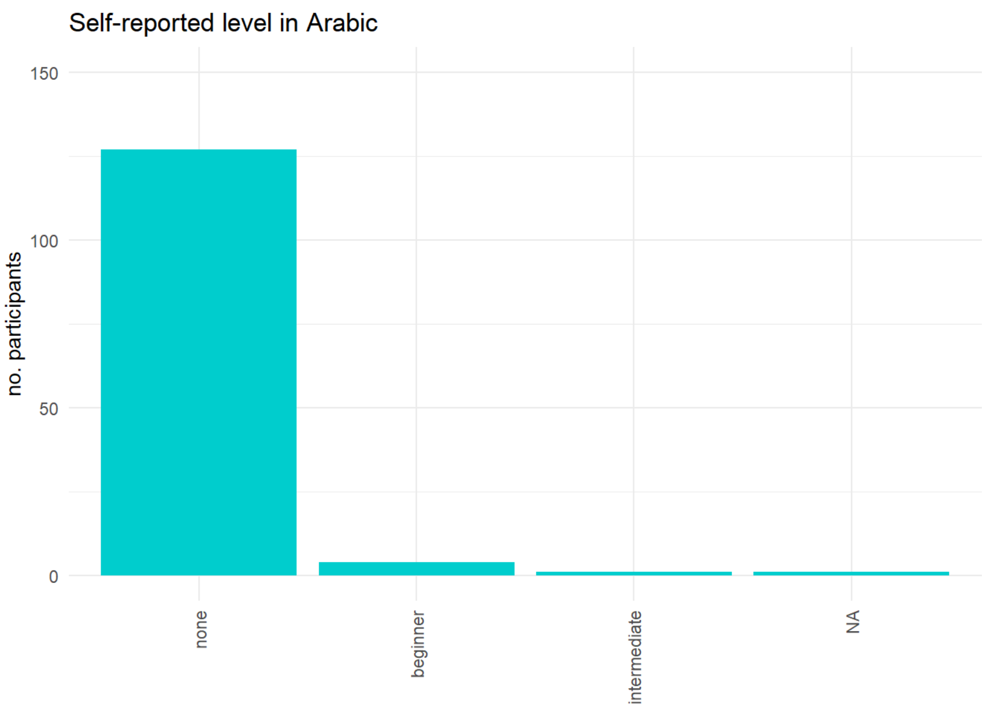


Figure V.12: L1 English participants' self-reported proficiency in L2 Arabic

Exploring self-reported ability further, the vast majority of English-speaking participants reported no ability across reading, writing, listening and speaking in Arabic.

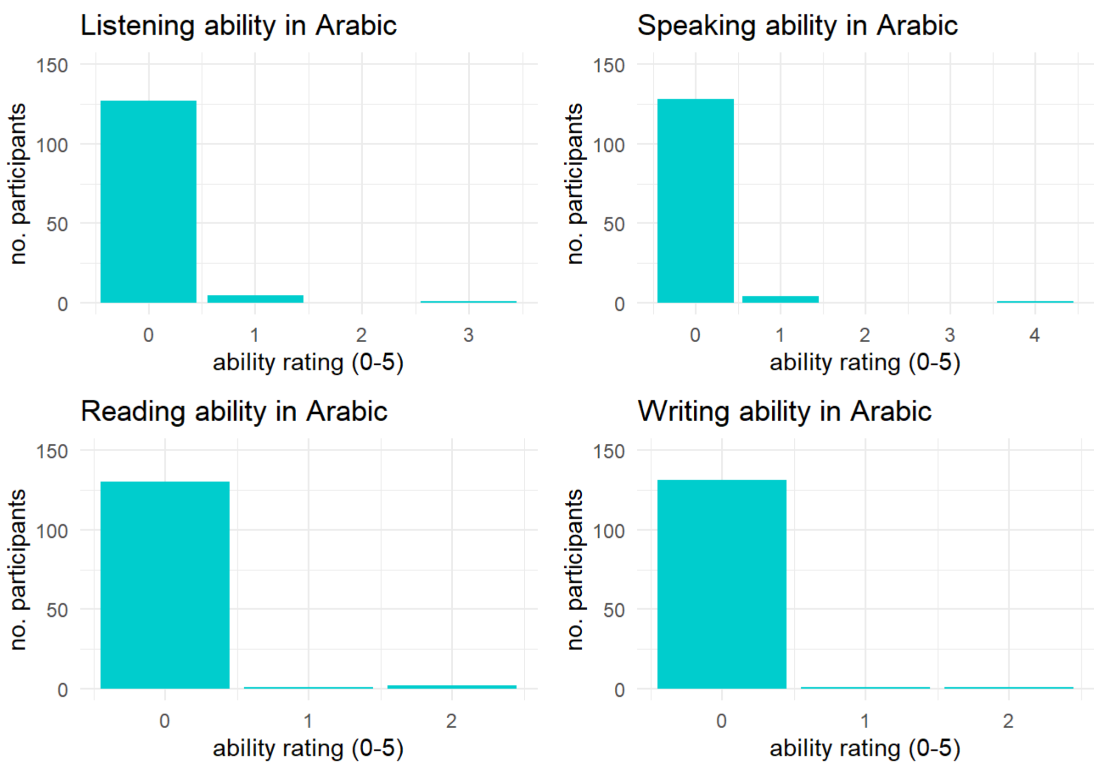


Figure V.13: L1 English participants' self-reported ability in different Arabic language skills

Taken together, these findings confirm that English participants had little to no experience with the Arabic language. Those participants who did have more experience with Arabic are further considered when discussing exclusions below.

Task set up data

The following descriptive analysis offers insight into the environment of participants when they were completing the study. The relevant variables include:

1. Equipment set up (audio and devices)
2. Distractions (reported amount and type)
3. Location (compared to nationality for migration insight)

This section draws on data from the audio-visual matching task, which was combined with demographic questionnaire responses. As some participants changed devices early on in the study for technical reasons, this data captures what device was being used for the main part of the experiment.

Participants were asked to use headphones or earphones, if possible, to complete the study. They were also recommended to use Chrome on a PC or Laptop to avoid technical issues. The figures below outline the audio equipment of participants, which was similar across both language groups, where the majority used their device speakers. Additionally, Most participants completed the study on a computer or mobile phone, with only a few people using tablets. Most English-speaking participants used a computer or laptop, whereas most Arabic-speaking participants used a mobile phone.

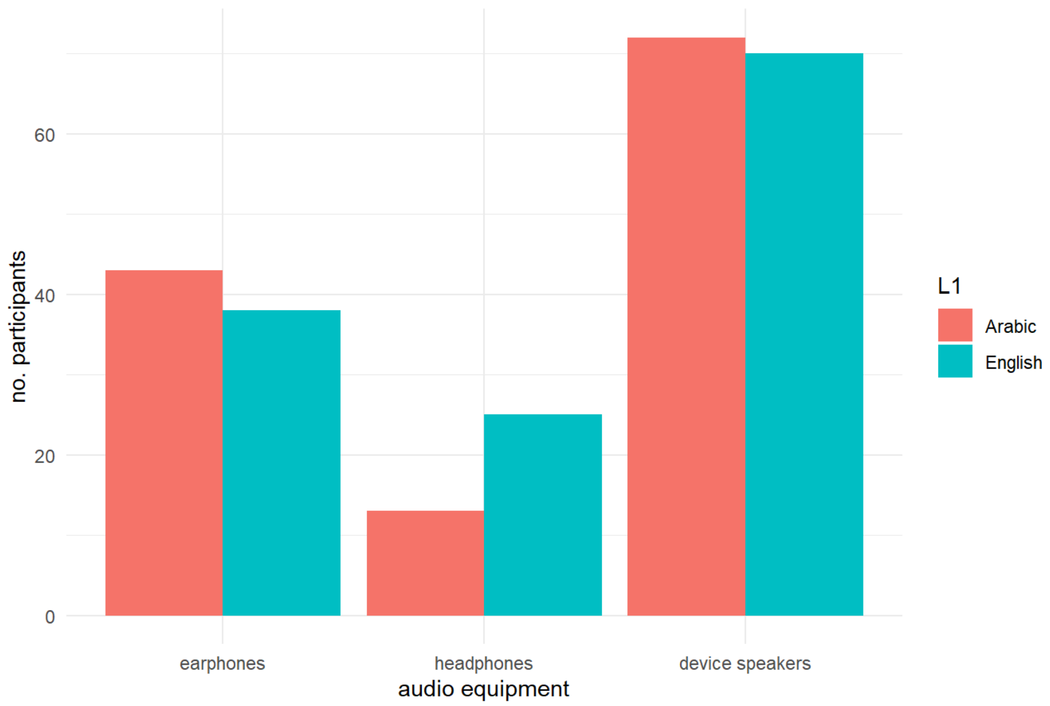


Figure V.14: Number of participants who used different types of audio equipment

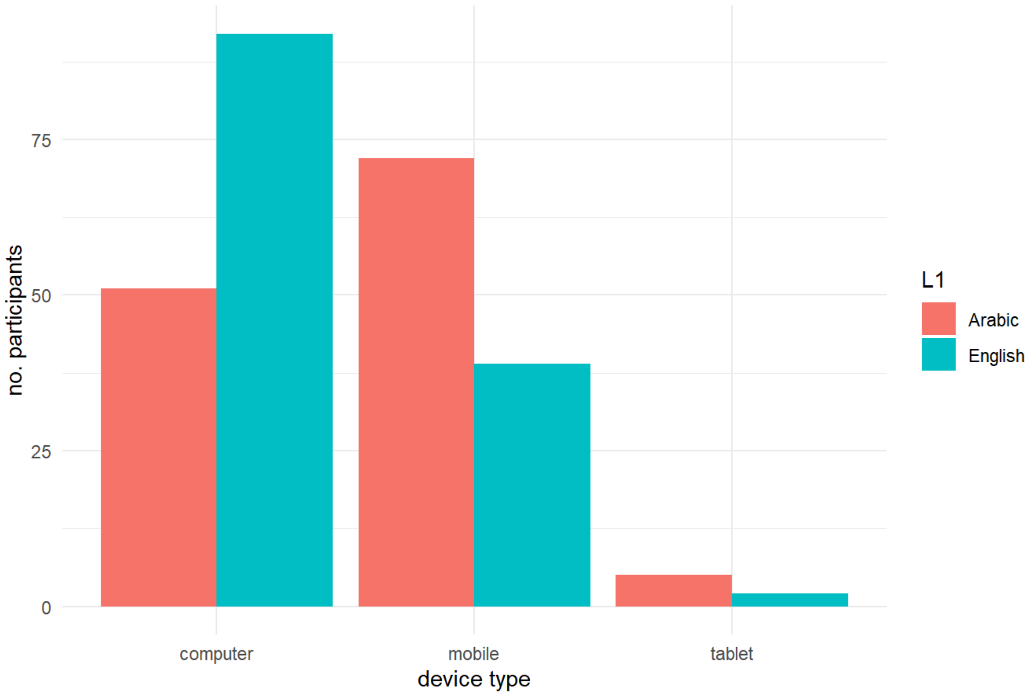


Figure V.15: Number of participants who used different devices to complete the study

To understand the extent to which participants were distracted by external factors, they were asked to use a sliding scale (0-100) to report their level of distraction during the study.

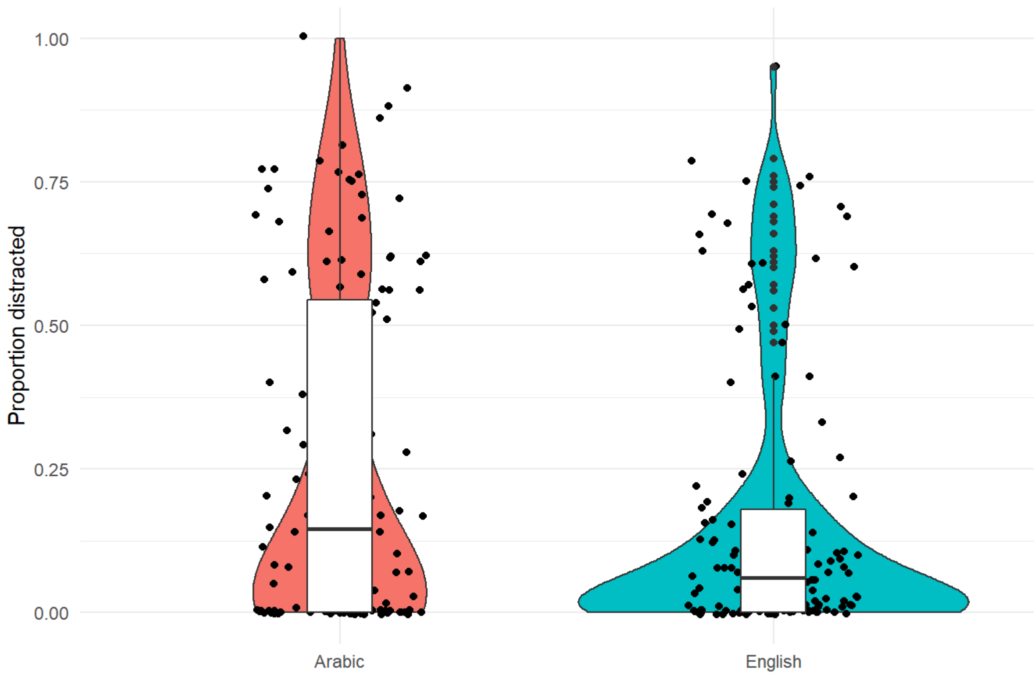


Figure V.16: Amount of reported distraction for both language groups

Participants were also asked what type of distractions they experienced and were able to select as many from the list as were relevant.

Figure V.17: Number of participants who reported each type of distraction during the study for both language groups

Some participants completed the study in a location that differed from their country of declared nationality. To see where people were when they completed the study, maps were generated for Arabic and English speakers separately below. Most Arabic-speaking participants either completed the study in Saudi Arabia or the UK. Aside from these locations most participants from Algeria, Egypt, Jordan and Iraq completed the study in their countries of declared nationality. Other locations of Arabic-speaker participation included Turkey, Australia, USA and the Netherlands.

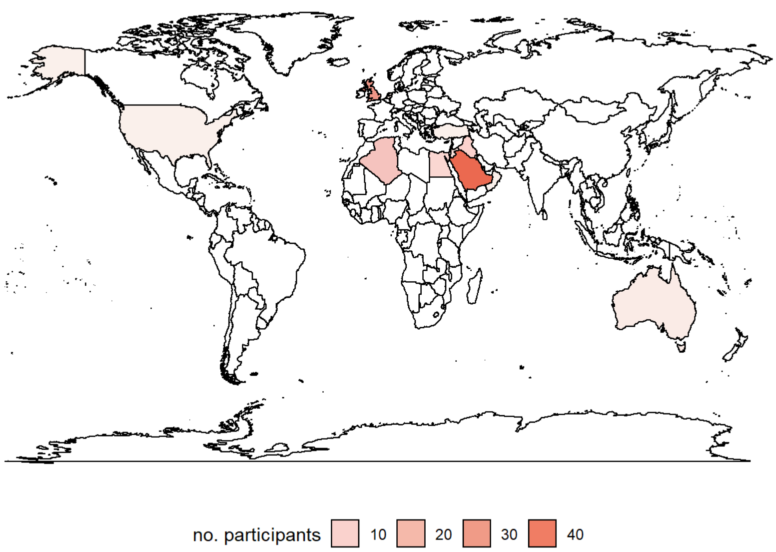


Figure V.18: Number of L1 Arabic participants in each country when completing the study

L1 English-speakers predominantly completed the study in the UK, with a number of participants spread across Europe, North America and Australia.

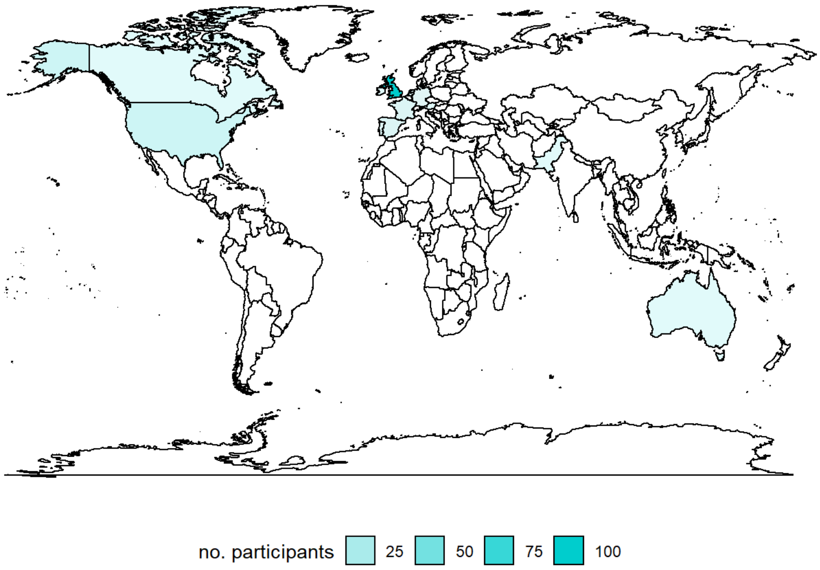


Figure V.19: Number of L1 English participants in each country when completing the study

In order to gain some speculative insight into migration, the table below shows that most participants completed the study in their country of reported nationality. More Arabic participants were abroad at the time of the study than English participants.

Table V.2: No. of participants who completed the study in a country that differed from their nationality

	L1 Arabic		L1 English	
	n	%	*n*	%
abroad	128		133	
… no	86	67.2%	104	78.2%
… yes	42	32.8%	29	21.8%

The figures below plot the location and nationalities of those who were abroad when completing the study.

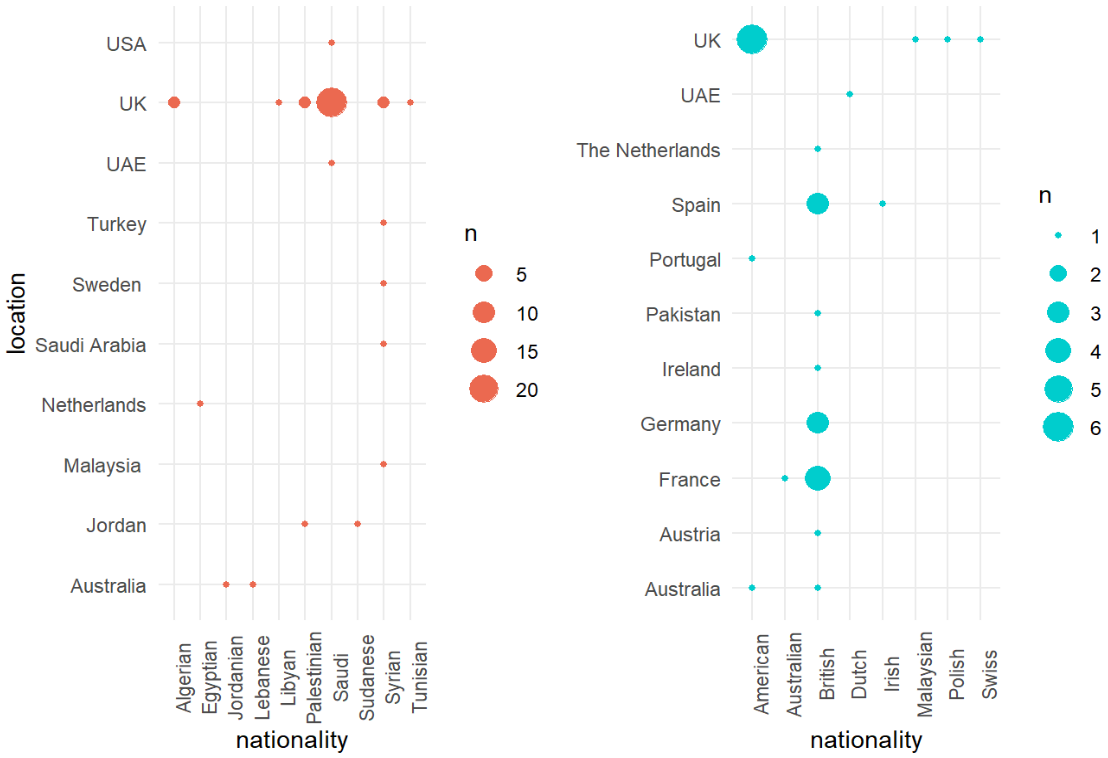


Figure V.20: Number of participants by location and nationality for both language groups

Task completion data

This section explores the timing data from the word learning phase and matching task, as data quality assessment. Most participants completed the learning and testing on the same day. The completion times were generally longer with more variation for the Arabic-speaking participants compared to the English-speaking participants for both learning and testing. The figures below explore this data further, demonstrating that both learning and testing exhibit extreme values for the Arabic speakers.

Table V.3: Overview of timing data to complete the two phases

	L1 Arabic			L1 English		
	n	**Mean**	**SD**	***n***	**Mean**	**SD**
Word learning time (mins)	128	14.503	19.41	133	10.38	2.025
Word matching time (mins)	128	5.586	4.211	133	4.34	1.064
Completed both on same day	124	96.9%		132	99.2%	

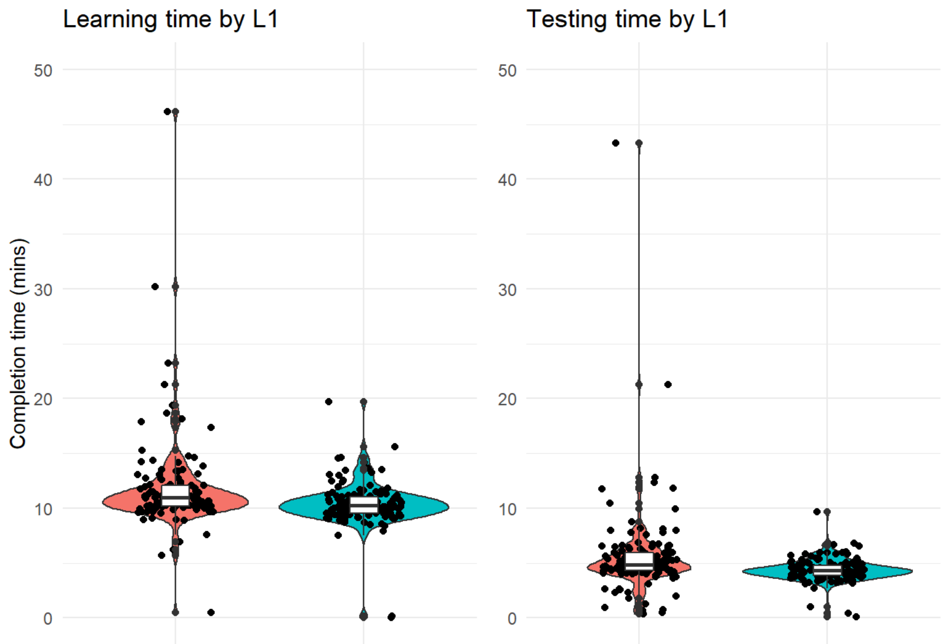


Figure V.21: Mean response time (RT) for both L1 groups to finish each phase

Outliers were calculated as exceeding the mean plus 2.5 standard deviations for each language group separately. Participants were also excluded if they did not complete the tasks on the same day. This results in the loss of data from three English-speakers and nine Arabic speakers. The figures below show that extreme values affected the timings for the matching task more than the learning phase, which again would be expected.

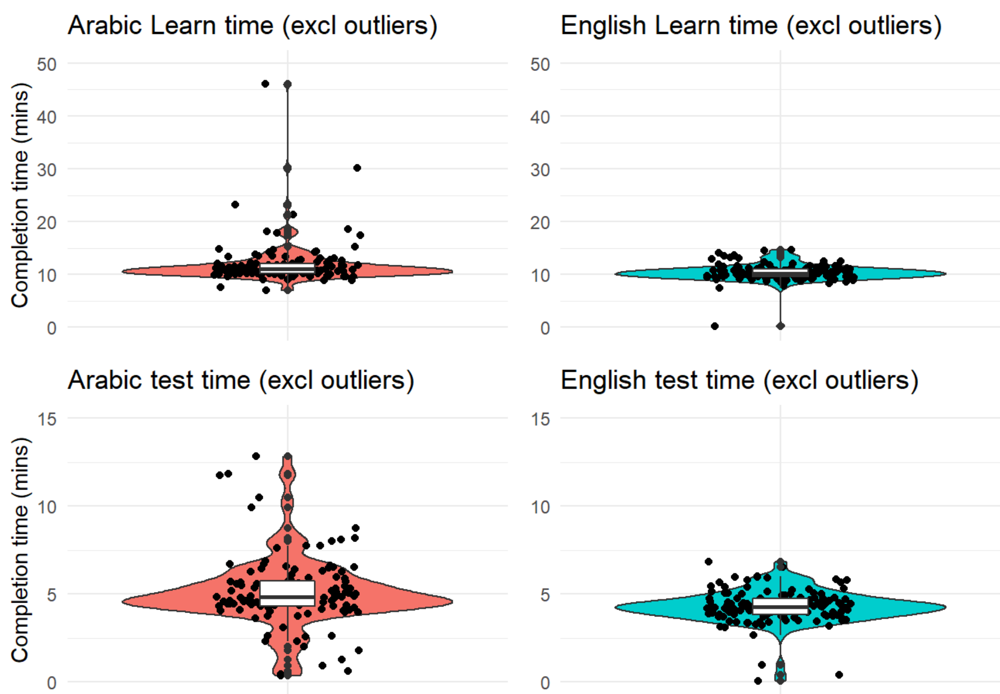


Figure V.22: Mean RT for both L1 groups completing each phase, with outliers excluded

There are some particularly low values, reflecting that timings reset in Gorilla if a participant leaves the study and then reopens or refreshes the browser. Excluding those that started and finished on different days is the only way to manage data quality for these participants.

Exclusions

This leads onto talking about exclusions more broadly and what data is used for the main analysis. To address issues of validity and improve the quality of the data, the variables reported below are used to inform necessary exclusions.

Age: The research question focuses on **adult** language acquisition and therefore should not include participants under the age of 16.

Vision/hearing: Participants that do not report normal or corrected to normal vision and hearing are excluded, as this is critical for an audio-visual matching task. Additionally, as the rate of hearing difficulty in older individuals (aka Presbycusis) increases in the population of adults over 65 years old (Gates & Mills, 2005), participants over this age are also excluded from the study.

Dyslexia/cognitive impairments: Participants are not excluded on this basis for two reasons, (1) There was clearly confusion on the interpretation of the question by Arabic-speaking participants, (2) dyslexia and other cognitive impairments are not widely assessed in many Arabic-speaking countries and there is more stigma attached, meaning self-report is unlikely to be a reliable indicator.

Arabic language ability: English participants were required to have little to no experience in Arabic, particularly in relation to literacy in the Arabic script. Therefore, those reporting reading ability above zero in Arabic are excluded.

Table V.4: Overview of data excluded for demographic reasons

	L1 Arabic			L1 English		
	n	**Mean**	**SD**	***n***	**Mean**	**SD**
Age	7	27	13.077	13	59.231	19.036
Vision/hearing impairment	5	71.4%		1	7.7%	
L2 reading ability	7	3.857	0.9	13	0.385	0.768

The primary concern for data quality in terms of task set up was amount of distraction. Stringent controls could have been placed on audio setup and screen size, but this was not deemed sufficiently beneficial to justify the number of participants it would exclude.

Distractions: Participants who reported disproportionate levels of distraction are excluded. This is calculated as those reporting a level of distraction greater than the mean + SD*2.5.

Table V.5: Amount of distraction reported by excluded participants

	L1 Arabic			L1 English		
	n	**Mean**	**SD**	***n***	**Mean**	**SD**
prop_distraction	3	0.93	0.062	1	0.95	NA

Same day completion: As this study was designed to teach and test new language on the same day, it was important that no participant data was included that involved a sleep cycle between teaching and testing. Therefore, any participants that completed the word matching task on a different day to when they began the word learning are excluded.

Completion times: It was decided not to exclude participants on the basis of extreme completion times, as long as they had completed the study on the same day. This was to take into account expected interruptions and pressures of participating outside of a lab, where participants had been told they could step away from the experiment and return later.

This was calculated above, showing that five participants would be excluded on this basis (four Arabic-speakers and one English-speaker)

Following the rationale laid out above, a new data set was created for the main analysis. After exclusions, the data from 114 L1 Arabic speakers (5478 observations) and 118 L1 English speakers (5664 observations) remained. A comparison of the original and new data set shows these exclusions resulted in 11% loss of both the English-speaker data and the Arabic-speaker data in the match task.

The same exclusion criteria were applied to the data from the word learning phase, which led to the same loss of 11% of data for both language groups. The same number of participants remained, with 5472 L1 Arabic observations and 5668 L1 English observations.

Appendix VI: Background questionnaire

Arabic version

[58]***Part 1: Demographic questions***

1. *How old are you? (open)*
2. *Which gender do you identify as? (multiple choice: male, female, not listed, prefer not to say)*
3. *What is your nationality? (open)*
4. *What country do you live in now? (open)*
5. *What is the highest level of education you have completed? (multiple choice: no schooling, primary school, secondary school, bachelor's degree, master's degree, doctorate, other professional qualification)*

Part 2: Arabic language questions

6. *Is Arabic your first or main language? (binary choice: yes, no)*
7. *What Arabic dialect do you speak? (open)*

Part 3: Language exposure questions

8. *On a typical weekday, what do you listen to in Dialect? (Radio/Podcast, Film/TV/Series, Social Media/Youtube, Music, Lecture/Lessons/Presentations, People around you, voice messages, nothing)*
9. *On a typical weekday, what do you listen to in MSA? (see previous)*
10. *On a typical weekday, what do you listen to in English? (see previous)*
11. *How many hours do you listen in each language on a typical day? (Dialect, MSA, English, Other: Sliding scale 0-12)*
12. *On a typical weekday, what do you read in Dialect? (social media, newspapers/magazines/articles, books, internet searches/information, personal writing/lists/notes, maps/signs, labels on food/medicine, nothing)*
13. *On a typical weekday, what do you read in MSA? (see previous)*
14. *On a typical weekday, what do you read in English? (see previous)*
15. *How many hours do you read in each language on a typical day? (Dialect, MSA, English, Other: Sliding scale 0-12)*
16. *On a typical weekday, when do you interact with others in Dialect? (neighbours/local community, friends/family, in a classroom, at work, on video calls, on phone calls, messaging via email/sms/social media, nothing)*
17. *On a typical weekday, when do you interact with others in MSA? (see previous)*
18. *On a typical weekday, when do you interact with others in English? (see previous)*
19. *How many hours do you interact in each language on a typical day? (Dialect, MSA, English, Other: Sliding scale 0-12)*

[58] English translation of Arabic questionnaire

Part 4: English language questions

20. *How old were you when you started learning English? (open)*
21. *How many years have you been studying English? (open)*
22. *Have you ever lived in an English speaking country? (binary choice: yes, no)*
23. *How long have you lived in an English-speaking country for? (multiple choice: never, less than 6 months, more than 6 months, more than 1 year, more than 5 years)*
24. *How would you rate your English proficiency? (Likert scale: 0-5 for listening, speaking, reading and writing)*
25. *What is your highest English language qualification? (open or 'No formal qualification')*

Part 5: Other languages

26. *What other languages do you use apart from Arabic and English?*

الجزء 1

1. كم عمرك؟
2. ما جنسك؟ (ذكر \ أنثى \ آخر \ أفضّل عدم القول)
3. ما جنسيتك؟
4. ماهي الدولة التي تعيش فيها الآن؟
5. ما هو أعلى مستوى تعليمي وصلت إليه؟ (لايوجد , الابتدائية , الثانوية , البكالوريوس , الماجستير , الدكتوراه , المؤهلات المهنية الأخرى)

الجزء 2

6. هل اللغة العربية لغتك الأم/الأولى؟ (نعم/لا)
7. ما هي لهجتك العربية؟

الجزء 3

8. خلال يومك الاعتيادي ماهي الاشياء التي تستمع لها بلهجتك؟ (راديو/بودكاست، تليفزيون/فيلم/مسلسل، منصات التواصل الاجتماعي/يوتيوب، دروس/محاضرات/عروض، أشخاص حولك، رسائل صوتية ، لا شيء)
9. خلال يومك الاعتيادي ما هي الأشياء التي تستمع لها باللغة العربية الفصحى؟
10. خلال يومك الاعتيادي ماهي الاشياء التي تستمع لها باللغة الانجليزية؟
11. كم ساعة تستمع بكل لغة؟ (لهجتك/ اللغة العربية الفصحى/ اللغة الانجليزية/ لغات اخرى)
12. خلال يومك الاعتيادي ماهي الاشياء التي تقرأها بلهجتك؟ (منصات التواصل الاجتماعي، جرائد/مجلات/ مقالات، كتب، بحث بالانترنت و معلومات، كتابات شخصية/ملاحظات/قوائم، خرائط و اشارات، ملصقات الاطعمة والادوية، لاشيء)
13. خلال يومك الاعتيادي ما هي الأشياء التي تقرأها باللغة العربية الفصحى؟
14. خلال يومك الاعتيادي ماهي الاشياء التي تقرأها باللغة الانجليزية؟
15. كم ساعة تقرأ بكل لغة؟ (لهجتك/ اللغة العربية الفصحى/ اللغة الانجليزية/ لغات أخرى)

16. خلال يومك الاعتيادي متى تتعامل مع الآخرين بلهجتك؟ (مع الجيران/المجتمع المحلي، مع الأصدقاء/العائلة ، في الفصول الدراسية، في العمل، في مكالمات الفيديو، في مكالمات الهاتف ، مراسلات عبر البريد الالكتروني/الرسائل النصية/منصات التواصل الاجتماعي، لا شيء)
17. خلال يومك الاعتيادي متى تتعامل مع الآخرين باللغة العربية الفصحى؟
18. خلال يومك الاعتيادي متى تتعامل مع الآخرين باللغة الانجليزية؟
19. كم ساعة تتعامل مع الآخرين بكل لغة؟ (لهجتك/ اللغة العربية الفصحى/ اللغة الانجليزية/ لغات أخرى)

الجزء 4

20. كم كان عمرك عندما بدأت تعلم اللغة الانجليزية ؟
21. كم سنة درست اللغة الانجليزية ؟
22. هل سبق وان عشت في مدينة لغتها الأولى الانجليزية ؟ (نعم، لا)
23. كم سنة عشت في مدينة لغتها الأولى الانجليزية ؟ (لا يوجد، اقل من ستة أشهر، أكثر من ستة أشهر، أكثر من سنة، أكثر من خمسة سنوات)
24. ما هو تقييمك لطلاقتك باللغة الانجليزية ؟ (التحدث، الاستماع، القراءة، الكتابة)
25. ماهي أعلى شهادة حصلت عليها باللغة الانجليزية؟

الجزء 5

26. ماهي اللغات الأخرى التي تستخدمها باستثناء اللغة العربية والانجليزية؟

English version

Part 1: Demographic questions

1. How old are you? (open)
2. Which gender do you identify as? (multiple choice: male, female, not listed, prefer not to say)
3. What is your nationality? (open)
4. What country do you live in now? (open)
5. What is the highest level of education you have completed? (multiple choice: no schooling, primary school, secondary school, bachelors degree, masters degree, doctorate, other professional qualification)

Part 2: Arabic language questions

6. Is English your first or main language? (binary choice: yes, no)
7. What variety of English do you speak? (open)

Part 3: Language exposure questions

8. On a typical weekday, what do you listen to in English? (Radio/Podcast, Film/TV/Series, Social Media/Youtube, Music, Lecture/Lessons/Presentations, People around you, voice messages, nothing)
9. On a typical weekday, what do you listen to in Arabic? (see previous)
10. On a typical weekday, what do you listen to in Other languages? (see previous)

11. How many hours do you listen in each language on a typical day? (English, Arabic, Other: Sliding scale 0-12)
12. On a typical weekday, what do you read in English? (social media, newspapers/magazines/articles, books, internet searches/information, personal writing/lists/notes, maps/signs, labels on food/medicine, nothing)
13. On a typical weekday, what do you read in Arabic? (see previous)
14. On a typical weekday, what do you read in other languages? (see previous)
15. How many hours do you read in each language on a typical day? (English, Arabic, Other, Other: Sliding scale 0-12)
16. On a typical weekday, when do you interact with others in English? (neighbours/local community, friends/family, in a classroom, at work, on video calls, on phone calls, messaging via email/sms/social media, nothing)
17. On a typical weekday, when do you interact with others in Arabic? (see previous)
18. On a typical weekday, when do you interact with others in Other languages? (see previous)
19. How many hours do you interact in each language on a typical day? (English, Arabic, Other, Other: Sliding scale 0-12)

Part 3: Arabic language questions

20. Have you ever lived in an Arabic speaking country? (binary choice: yes, no)
21. How long have you lived in an Arabic-speaking country for? (multiple choice: never, less than 6 months, more than 6 months, more than 1 year, more than 5 years)
22. How would you rate your Arabic proficiency? (Likert scale: 0-5 for listening, speaking, reading and writing)

Part 4: Other languages

23. What other languages do you use, apart from English and Arabic?

Appendix VII: Post-test questionnaire

English version

Part 1: Influence of written input

1. When learning the new words, when did you notice the English or Arabic spelling? (Near the start, near the middle, near the end, I didn't notice they were there)
2. Seeing the word written in Arabic letters…(made learning easier, made learning more difficult, made no difference, I didn't notice)
3. Seeing the word written in English letters… (see previous)
4. Hearing the word without the spelling…(see previous)
5. Did it make a difference seeing the written words? In what way?
6. I would have preferred to hear the words and see all of them…(written in Arabic, written in English, without the spelling)
7. In general, when learning new words, is it important to see the spelling the first time you hear it? (yes, no)
8. Why? (open)

Part 2: Speculative spelling

How would you spell the word for this picture?

9. 
10. 
11. 
12. 

Part 3: Language learning strategies

13. How did you try to learn or remember the words?

Read the statements and decide how true they are for you in relation to your experience learning the words today. Do you think that it was **always true for you**, **almost always true for you**, **sometimes true for you**, **almost never true for you**, or **never true for you**?

14. I created association with words or things I already knew. (always, almost always, sometimes, almost never, never)
15. I put the word in a context to remember it (e.g. a sentence, a story, a rhyme).

16. I made a mental image or imagined additional connections to help me remember.
17. I visualised the spelling of the word in Arabic in my mind.
18. I visualised the spelling of the words in English in my mind.
19. I repeated the words out loud or in my head.
20. I thought about or practised mouth positions to pronounce the words.
21. I thought about an action or movement to help remember the words.
22. I broke words down into syllables and sounds.
23. I tried to find patterns in the new words and sounds.
24. I tried to connect sounds and letters.
25. I tried to remember the first or last sounds.
26. I looked for similarities and contrasts between the pronunciation of the new words and the languages I know.
27. I grouped similar sounding words.
28. I used the English spelling to distinguish between similar sounds.
29. I used the Arabic spelling to distinguish between similar sounds.
30. I purposefully ignored the English spelling.
31. I purposefully ignored the Arabic spelling.
32. I found ways to test my memory and recall the new words.
33. I thought about my progress in learning the new words.
34. I noticed my mistakes and used that information to help me do better.
35. I noticed that I felt more relaxed or confident when I saw the Arabic spelling.
36. I noticed that I felt more relaxed or confident when I saw the English spelling.

Arabic version

الجزء 1

1. عندما تتعلم كلمات جديدة، متى تلاحظ الإملاء بين اللغة العربية واللغة الانجليزية؟ (عند البداية ، في المنتصف، عند النهاية. لا الاحظ ذلك)
2. مشاهدة الكلمات مكتوبة باللغة العربية...(يجعل التعلم اسهل، يجعل التعلم أصعب, لايوجد اختلاف. لم الاحظ الإملاء)
3. مشاهدة الكلمات مكتوبة باللغة الانجليزية...
4. الاستماع للكلمات من دون ملاحظة الإملاء...
5. هل أحدث ذلك فرقا في رؤية الكلمات المكتوبة؟ إلي أي مدي؟
6. افضّل ان استمع للكلمات واشاهدها...(مكتوبة باللغة العربية، مكتوبة باللغة الإنجليزية، من دون املاء)
7. بشكل عام، عندما تتعلم كلمات جديدة ، هل من المهم عندك أن ترى املاء الكلمة عند الاستماع لها لأول مرة ؟ (نعم ، لا)
8. لماذا ؟

الجزء 2

كيف تكتب املاء الكلمة بالصورة؟

9.

.10 

.11 

.12 

الجزء 3

13. كيف حاولت تعلم أو تذكر الكلمات ؟

اقرأ الجمل وقرر مدى صحتها بالنسبة لتجربتك في تعلم الكلمات الجديدة اليوم. هل تعتقد بأن التجربة كانت **دائماً صحيحه، غالباً صحيحه، صحيحه في بعض الاوقات، لم تكن صحيحة في الغالب، لم تكن صحيحة دائماً**.

14. لقد قمت بإنشاء روابط بكلمات و أشياء أعرفها من قبل. (دائما, دائما تقريبا, أحيانا, غالبا لا, أبدا)
15. لقد قمت بوضع الكلمة في سياق ما لكي أتذكرها (علي سبيل المثال، جملة، قصة، قافية).
16. لقد صنعت صورة تخيلية فى ذهني أو تخيلت روابط إضافية لمساعدتي على التذكر.
17. لقد تخيلت نطق الكلمات باللغة العربية في ذهني.
18. لقد تخيلت نطق الكلمات باللغة الإنجليزية في ذهني.
19. لقد قمت بتكرار الكلمات بصوت عالي او في ذهني.
20. لقد قمت بمحاولة تحريك فمي لنطق الكلمات.
21. لقد فكرت في فعل أو حركة لمساعدتي على تذكر الكلمات.
22. لقد قسمت الكلمات لمقاطع صوتية و اصوات.
23. لقد حاولت إيجاد أنماطاً في الكلمات والأصوات الجديدة.
24. لقد حاولت ربط الأصوات بالحروف.
25. لقد حاولت تذكر الأصوات الأولي أو الأخيرة.
26. لقد بحثت عن التشابهات والاختلافات بين نطق الكلمات الجديدة و نطقها في اللغات التي اعرفها.
27. لقد جمعت الكلمات ذات الأصوات المتشابهة.
28. لقد قمت بتهجئة الحروف باللغة الإنجليزية للتمييز بين الأصوات المتشابهة.
29. لقد قمت بتهجئة الحروف باللغة العربية للتمييز بين الأصوات المتشابهة.
30. لقد تجاهلت عن عمد تهجئة الحروف باللغة الإنجليزية.
31. لقد تجاهلت عن عمد تهجئة الحروف باللغة العربية.
32. لقد وجدت طرقاً لإختبار ذاكرتي و مدى تذكري الكلمات الجديدة.
33. لقد فكرت في مدى تقدمي في تعلم الكلمات الجديدة.
34. لقد لاحظت أخطائي و أستخدمت ما تعلمته لكي أتحسن.
35. لقد لاحظت أنني كنت أكثر راحة و ثقة حينما رأيت تهجئة الحروف باللغة العربية.
36. لقد لاحظت أنني كنت أكثر راحة و ثقة حينما رأيت تهجئة الحروف باللغة الإنجليزية.

Appendix VIII: Matching task data exploration Rmd output

The tables and figures below offer an overview of variables related to participant set up and environment when completing the study, as well as background information. The variables are plotted by accuracy and response time (RT) to inform inferential analysis.

Environmental factors

To explore the extent to which participant performance was influenced by environmental factors during completion of the study session. Frequencies are plotted by:

1. Audio setup (headphones, earphones, device speakers)
2. Device type (computer/laptop, mobile phone, tablet)
3. Distractions (reported amount and type)

Mean accuracy and RT by speaker is calculated to generate subset data frames and frequency plots. This is done by filtering correct answers and then creating a new data frame which summarises (1) the number of correct answers and (2) mean RT for the correct response of each participant, grouped by the chosen environmental variables.

The figure below plots the number of correct responses and mean RT by audio set up. The central tendency and distribution of scores and times are comparable, whether or not participants used headphones.

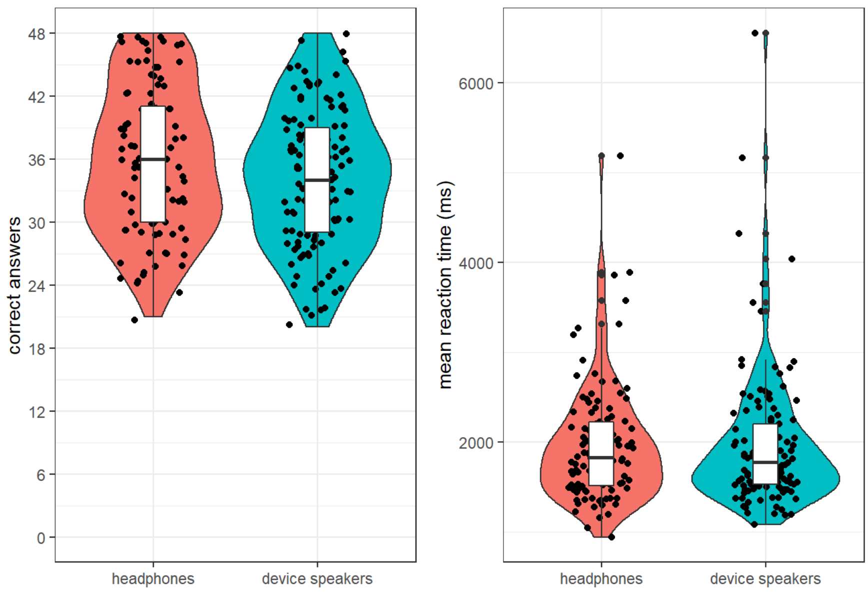


Figure VIII.1: Participant raw accuracy and RT by audio equipment

To assess whether audio setup influenced performance with each of the target contrasts, another data frame is created which summarises responses by target contrast. The violin plots below indicate marginally fewer correct answers when using device speakers for the difficult /f-v/ contrast words. Mean RT does not differ by target contrast or audio equipment.

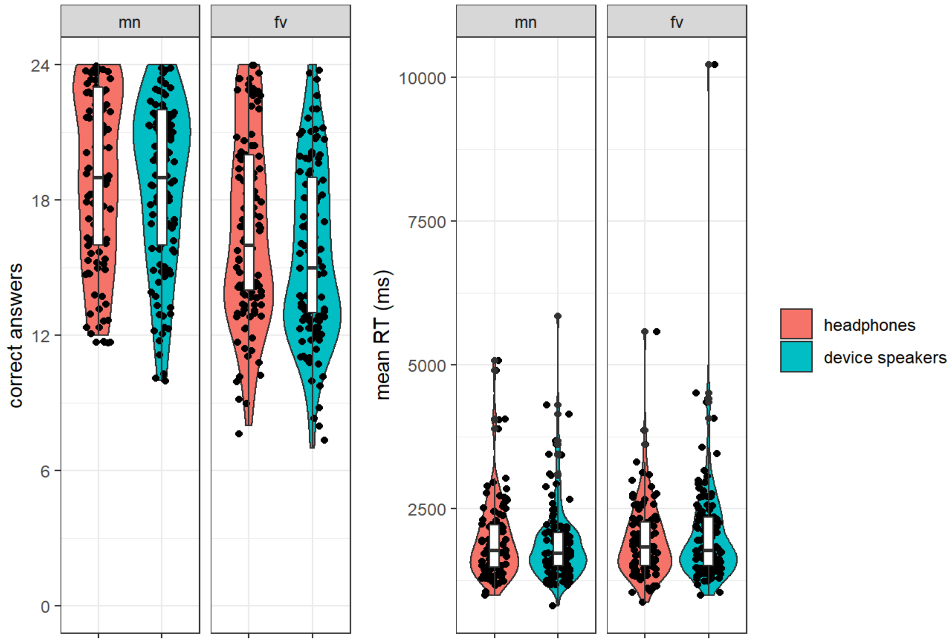


Figure VIII.2: Participant raw accuracy and RT by audio equipment and contrast

Violin plots also provide insight into participants' devices and whether that influences accuracy and RT. Performance is broadly comparable across devices, although accuracy and RT are marginally lower and faster, respectively, with mobile phones. The unusual shape of the plot for tablet users reflects the small sample.

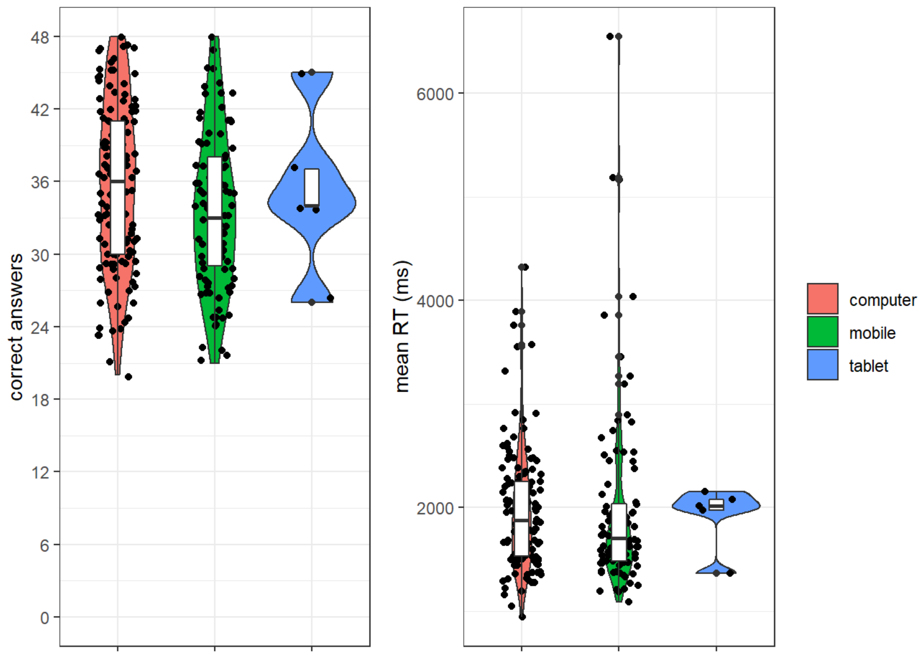


Figure VIII.3: Participant raw correct responses and RT by device type

The marginally lower scores with mobile phones could be for a number of reasons. Mobile phone screens are smaller than both computers and tablets, which could increase the difficulty of the task. Additionally, distractions may be more likely, such as interruptions from device notifications, as well as increased likelihood of multitasking and background distractions. It may also be that touchscreen response on mobile phones and tables may be marginally faster compared to a mouse/mousepad response.

The figures below plot L1 Arabic and English participants' reported proportion of distraction while completing the study by their raw accuracy and RT. Most participants indicate little to no distraction while completing the study. The trend line indicates a negative correlation between correct answers and amount of distraction, and positive correlation with RT, for Arabic-speakers but not English-speakers. Thus, as distraction increases, accuracy reduces and reaction time slows.

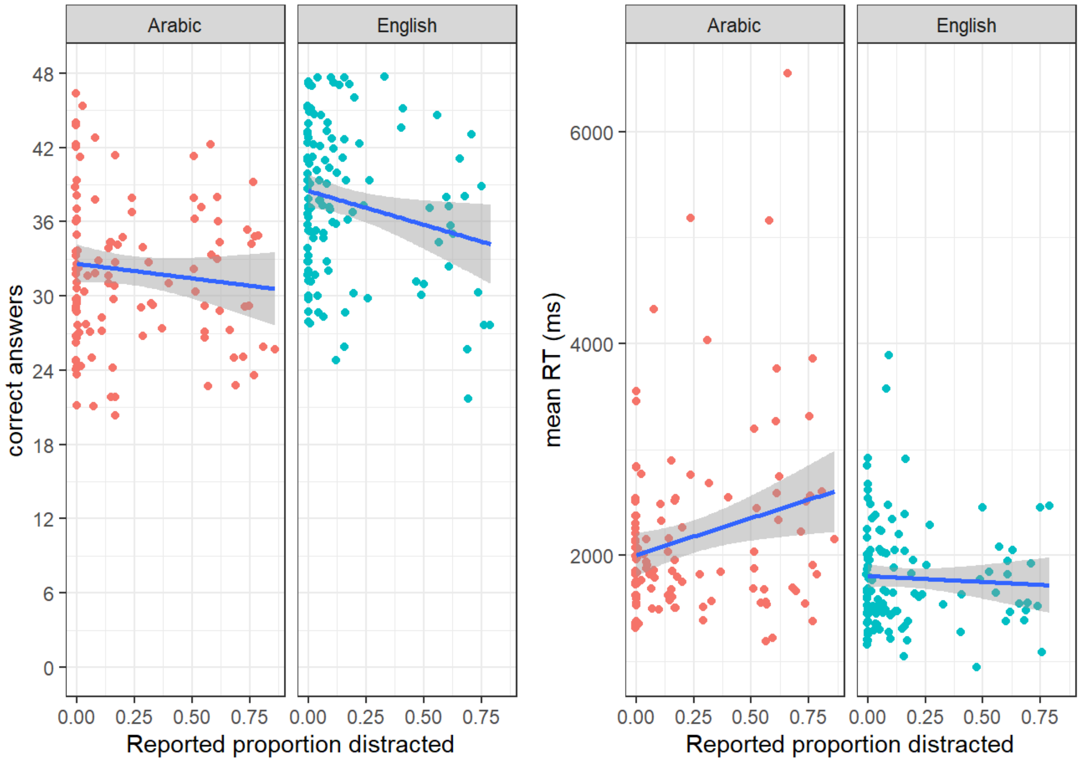


Figure VIII.4: Participant raw correct responses and RT by reported amount of distraction and L1 language group

Further insight is offered by exploring the types of reported distractions. As participants are able to select more than one distraction, new data frames are created for each of the distraction options offered and plotted independently. The plots below also show that most participants report no distraction during the study. The most common distractions reported are background noise, interruptions from a member of someone's household, and general tiredness. The figures below plot the number of correct answers and RT for participants by type of distraction, where the colours from previous plots represent the language groups.

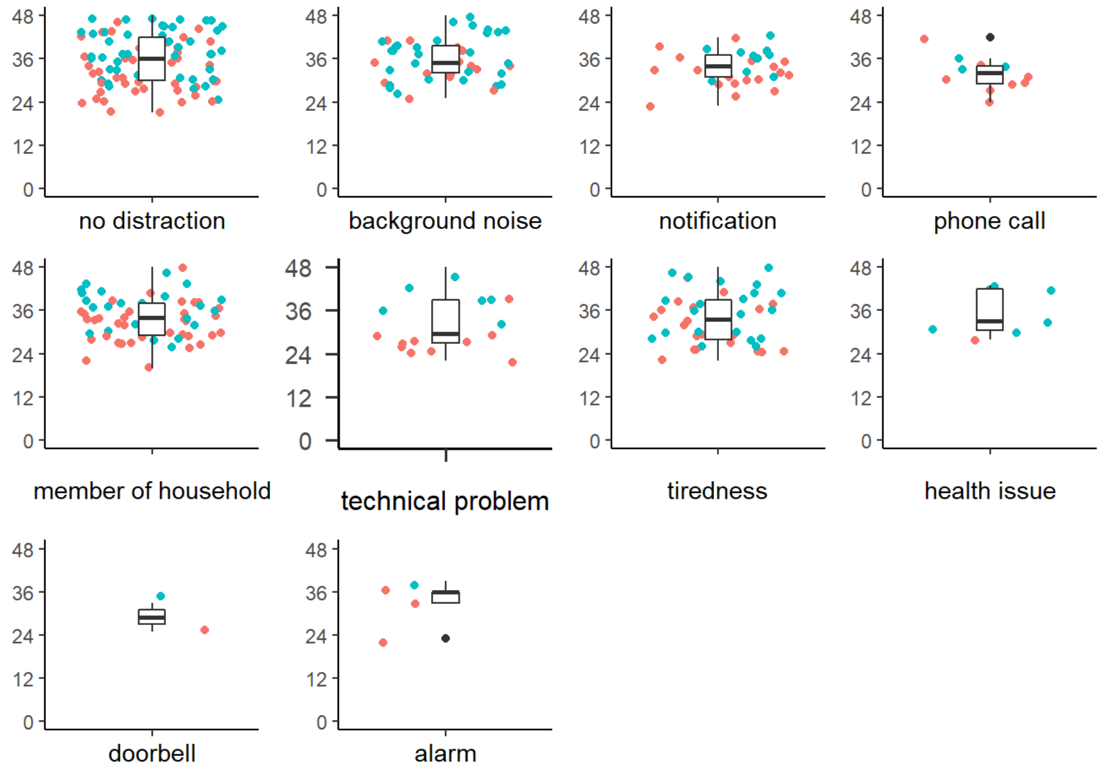


Figure VIII.5: Participant raw accuracy by type of distraction and L1 language group

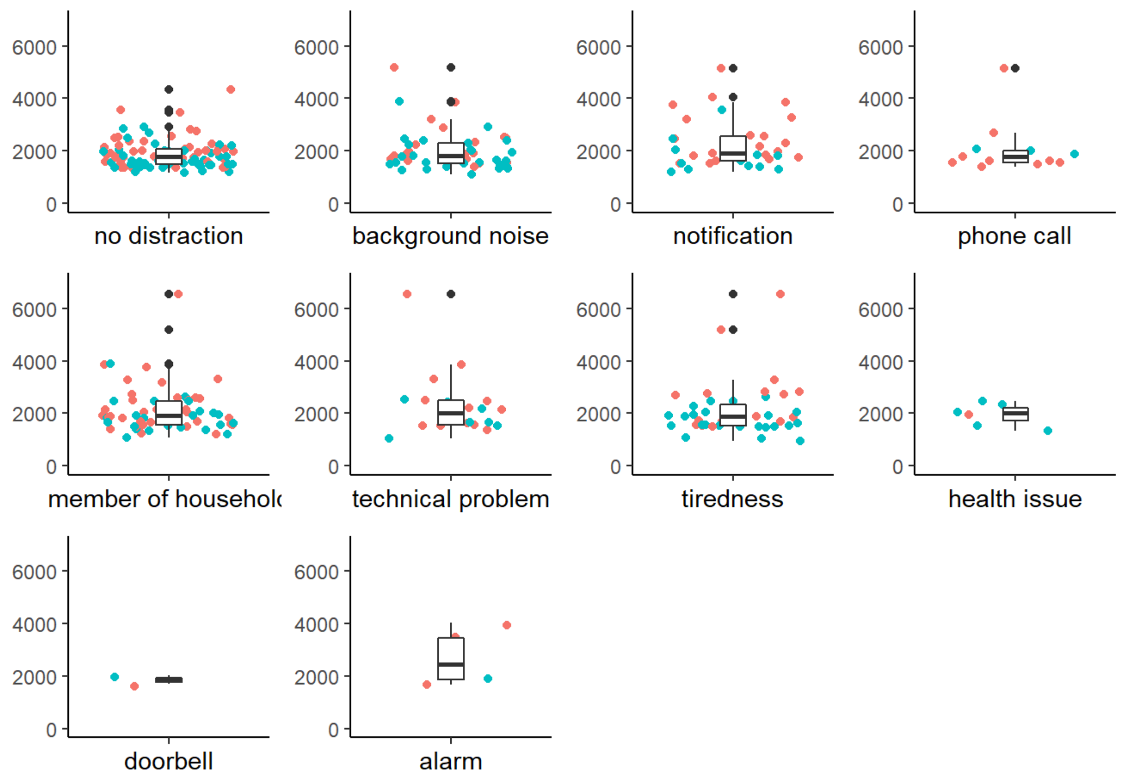


Figure VIII.6: Participant mean RT by type of distraction and L1 language group

Demographic factors

Next, demographic factors are explored and in relation to participant performance in the audio-visual matching task. Frequencies are plotted by:

1. first language
2. age
3. gender
4. level of completed education
5. dialect
6. nationality
7. location
8. other languages

As above, raw accuracy scores and mean RT by speaker are calculated to generate subset data frames and frequency plots, which summarise the number of correct answers of each participant by demographic variables.

It is predicted that L1 English-speakers would outperform L1 Arabic-speakers on the matching task, as the words are English pseudowords, compliant with the phonotactics and sound-spelling correspondences of English words. Additionally, all sound contrasts are well-established in English phonology. The violin plots below confirm this prediction, as several English participants perform at ceiling and the central tendency is higher than Arabic-speaking participants. L1 English participants also have faster mean RT. The rest of the descriptive analysis focuses on these groups separately, in relation to their task accuracy and reaction times.

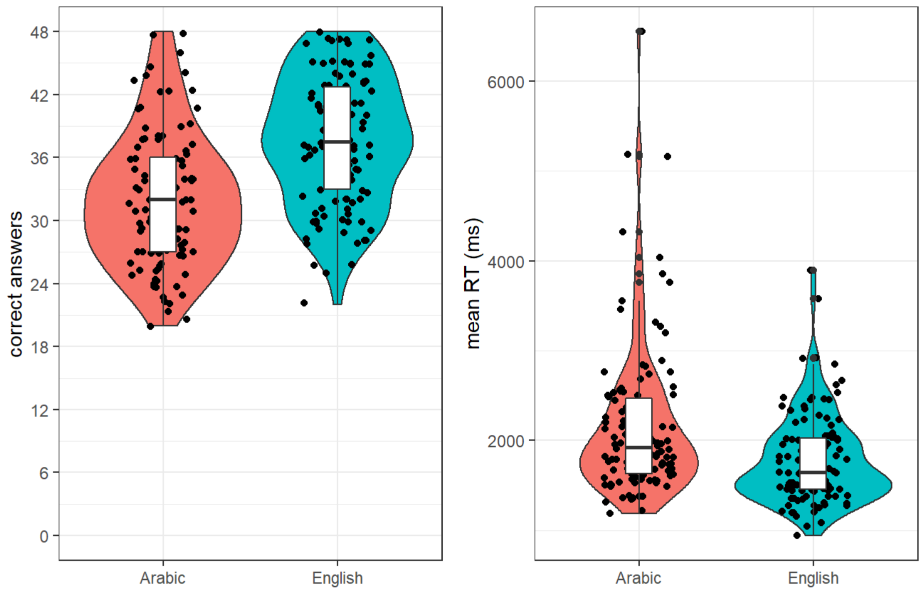


Figure VIII.7: Participant raw accuracy and mean RT by L1 language group

The age range is larger for English-speakers than Arabic-speakers. There is also a negative correlation between age and accuracy and positive correlation between age and RT for the English-speakers. Thus, older participants are less accurate and slower than younger participants, likely related to working memory and auditory sensitivity demands of the task.

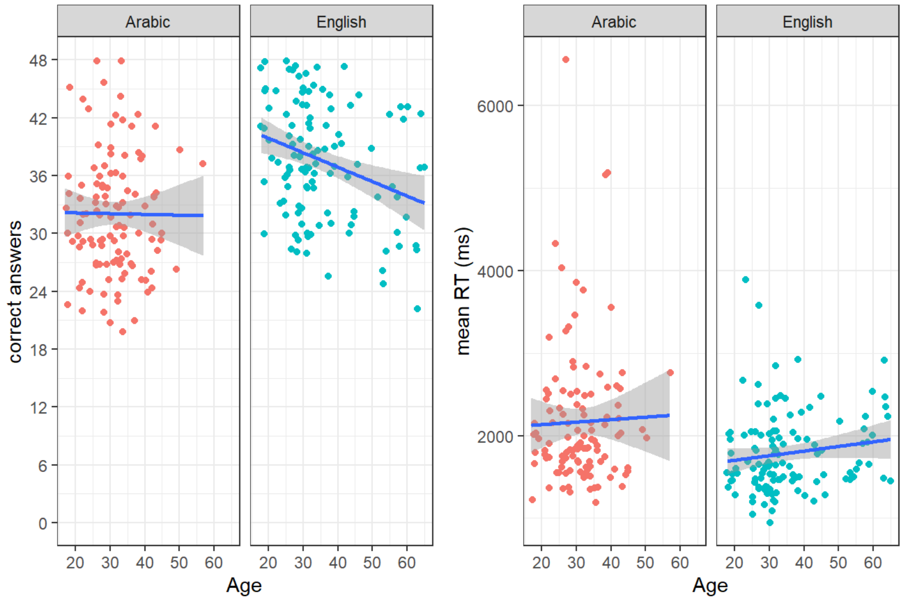


Figure VIII.8: Participant raw accuracy and mean RT by age and L1

Performance is broadly comparable across male (m) and female (f) participants in both language groups. Participants who do not identify as either male or female in the English-speaking group (n) demonstrate high performance; however, the small sample does not allow for further interpretation.

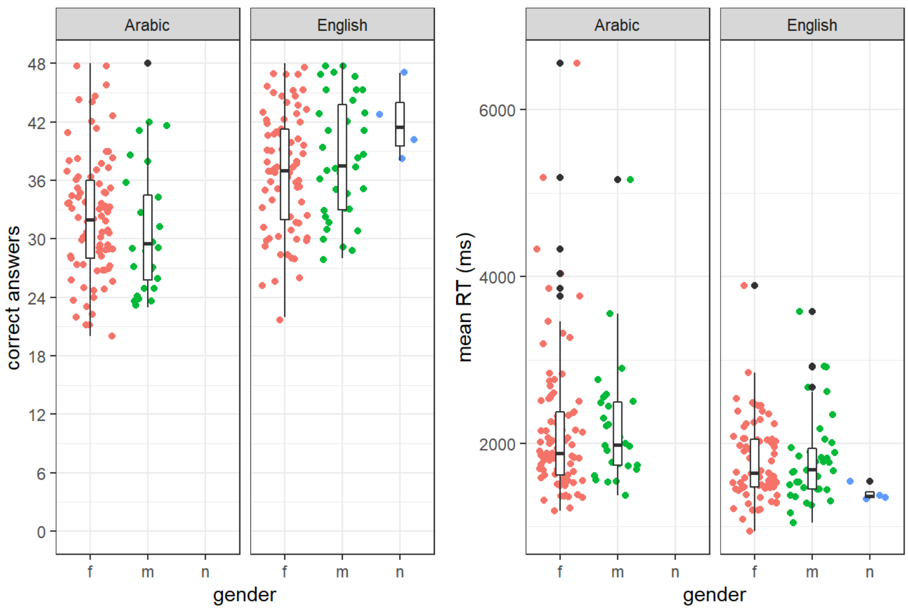


Figure VIII.9: Participant raw accuracy and mean RT by gender and L1

L1 English-speakers outperform L1 Arabic-speakers, regardless of educational background. Accuracy and RT across different Arabic-speaker educational levels appear comparable, especially for those with a university degree. Those with a maximum of secondary school education exhibit the largest spread of correct answers. The central tendency of accuracy for English-speakers appears to marginally improve between those with other professional qualifications and undergraduate degrees compared to those with postgraduate degrees.

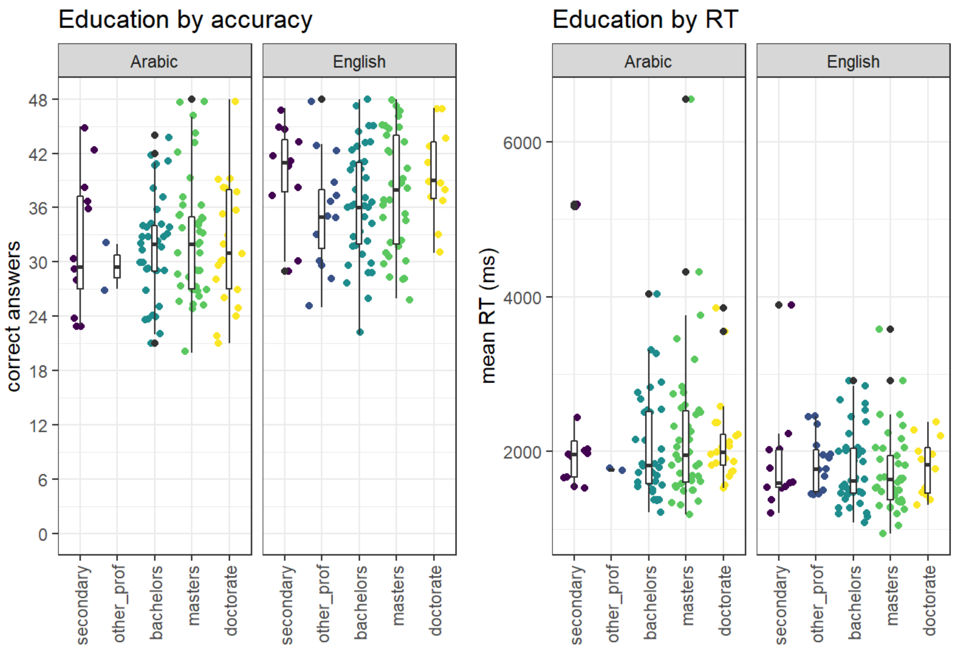


Figure VIII.10: Participant raw accuracy and mean RT by level of education and L1

As mentioned in appendix V, Arabic-speaker dialects are grouped into five regional varieties: Maghrebi, Egypto-Sudanic, Mesopotamian, Levantine, Gulf/ Peninsula. Dialect is then explored in relation to accuracy and RT in the matching task.

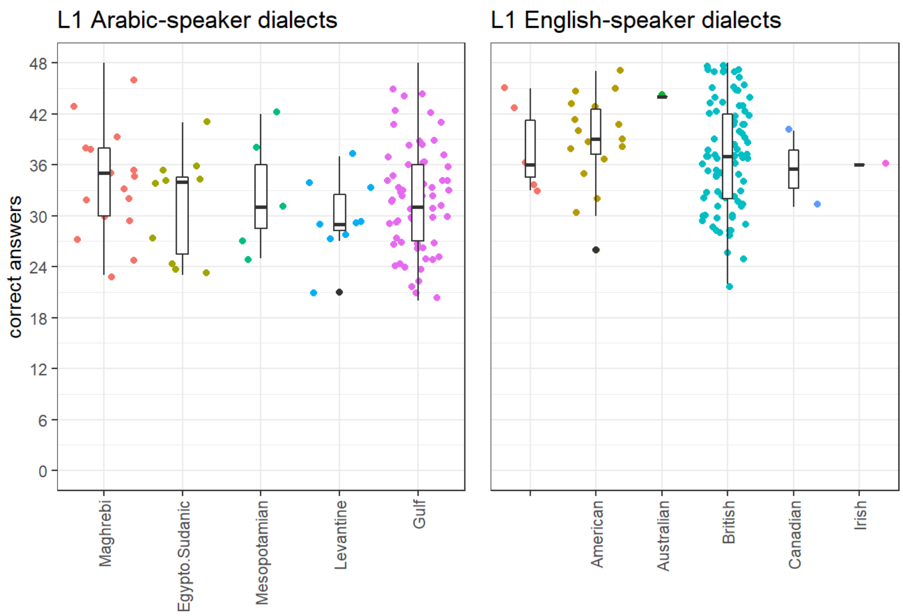


Figure VIII.11: Participant raw accuracy by dialect and L1

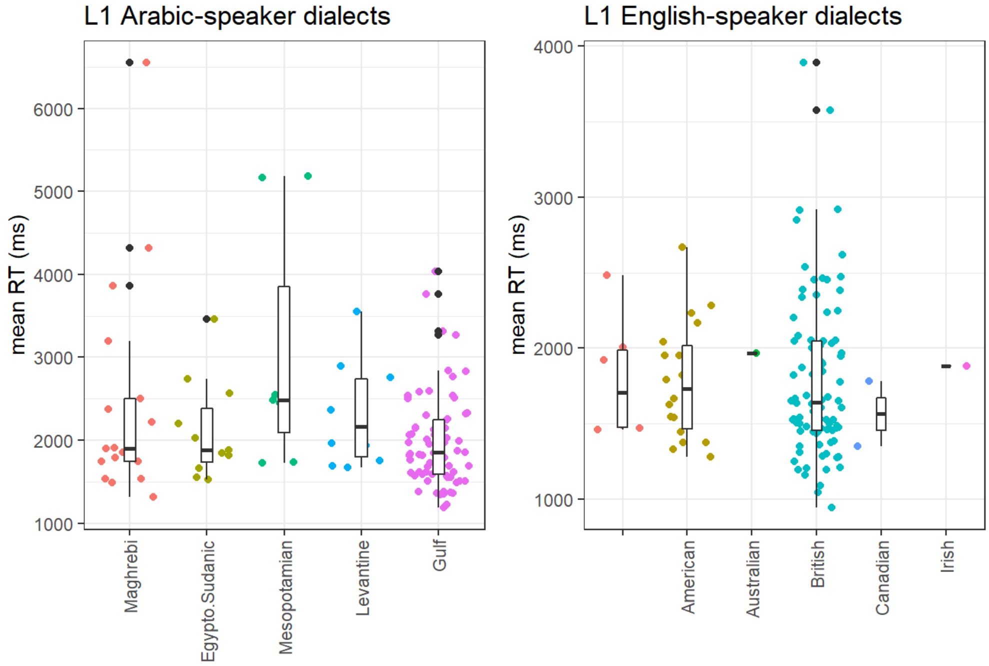


Figure VIII.12: Participant mean RT by dialect and L1

To ascertain whether certain dialect backgrounds have an advantage with the contrast /f-v/, accuracy is plotted by dialect and experimental contrast. This is of particular interest for the Maghrebi Arabic-speakers, whose dialect is more influenced by Romance languages that contain the /f-v/ contrast, such as French. The figure below indicates that this may have been the case. However, a more comparative sample is required to better assess this.

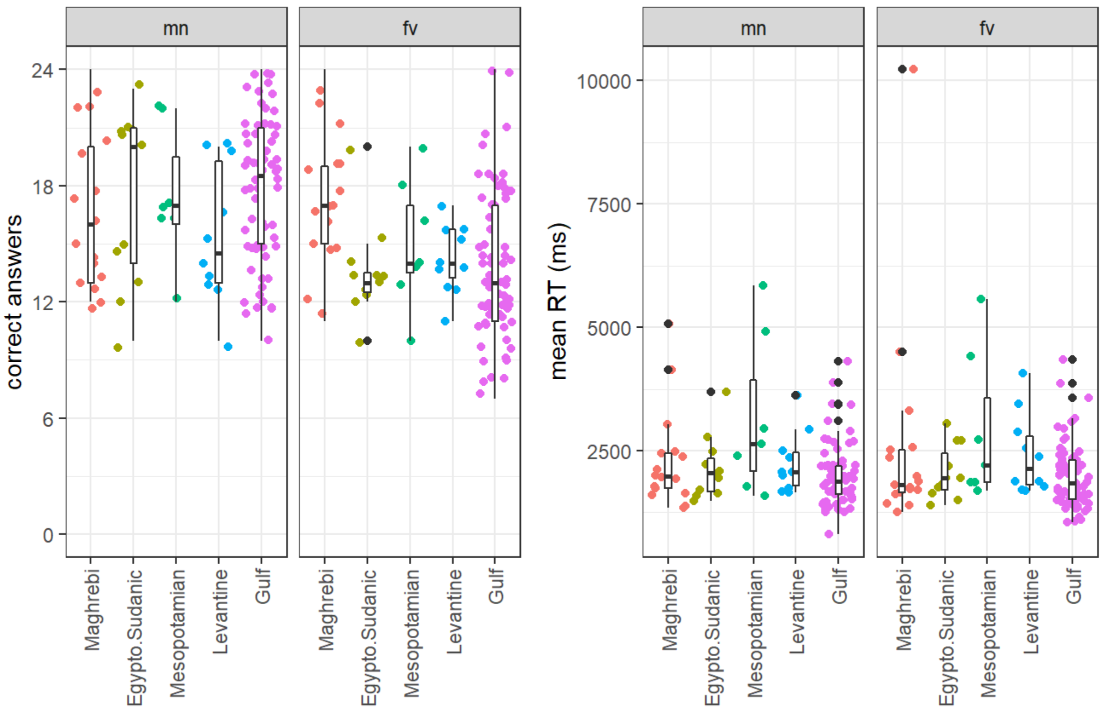


Figure VIII.13: L1 Arabic raw accuracy and mean RT by dialect and phonological contrast

The figures below plot raw accuracy and mean RT by nationality and location of both Arabic-speakers (AS) and English-speakers (ES). There is no clear advantage for nationality or location.

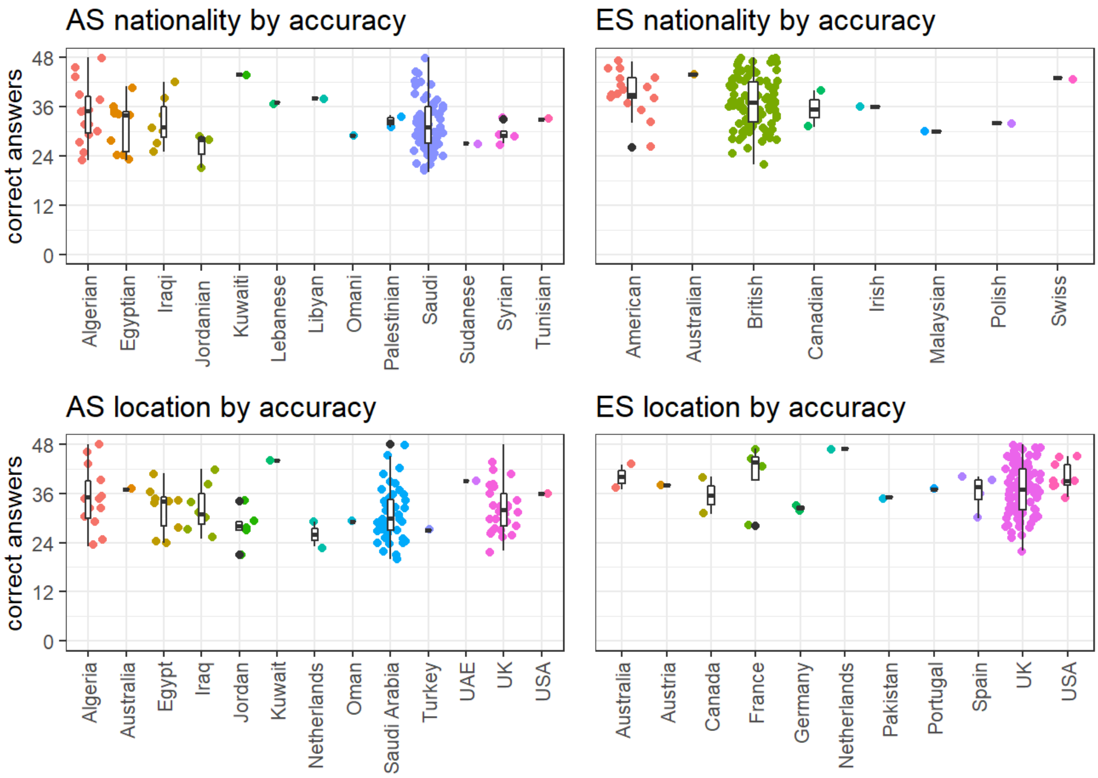


Figure VIII.14: Participant raw accuracy by nationality and location and L1

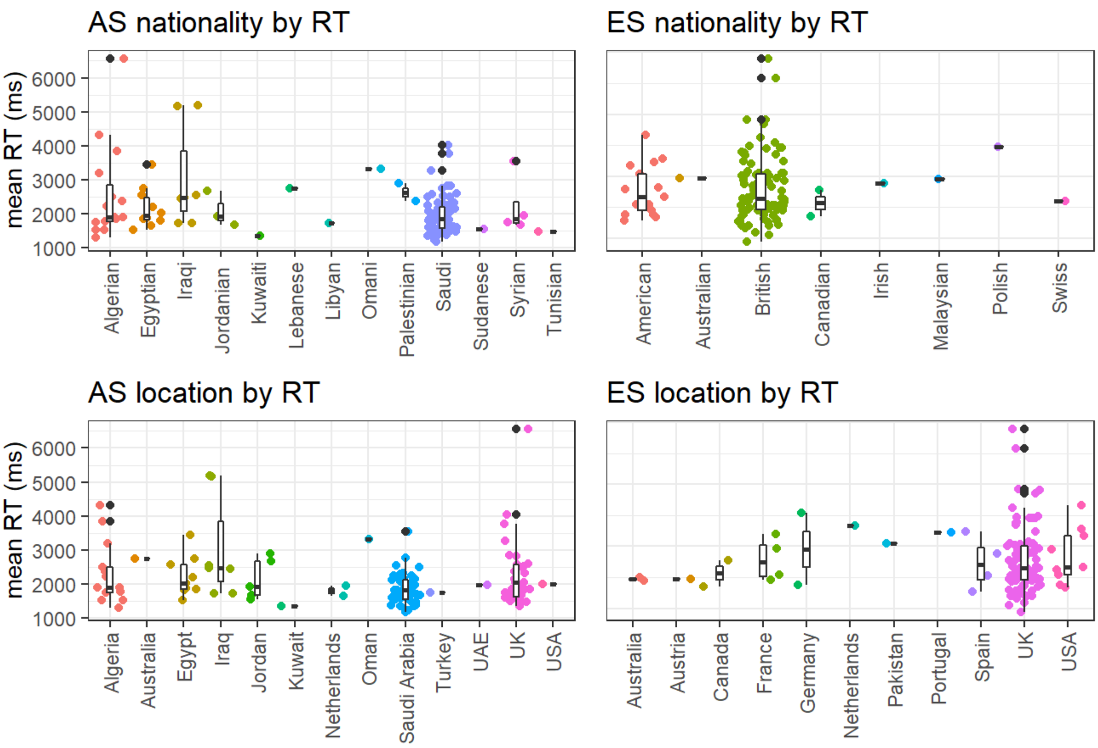


Figure VIII.15: Participant mean RT by nationality and location and L1

The figures below plot participant raw accuracy and mean RT by other languages that they know. There appears to be no clear advantage based on the languages reported by the participants

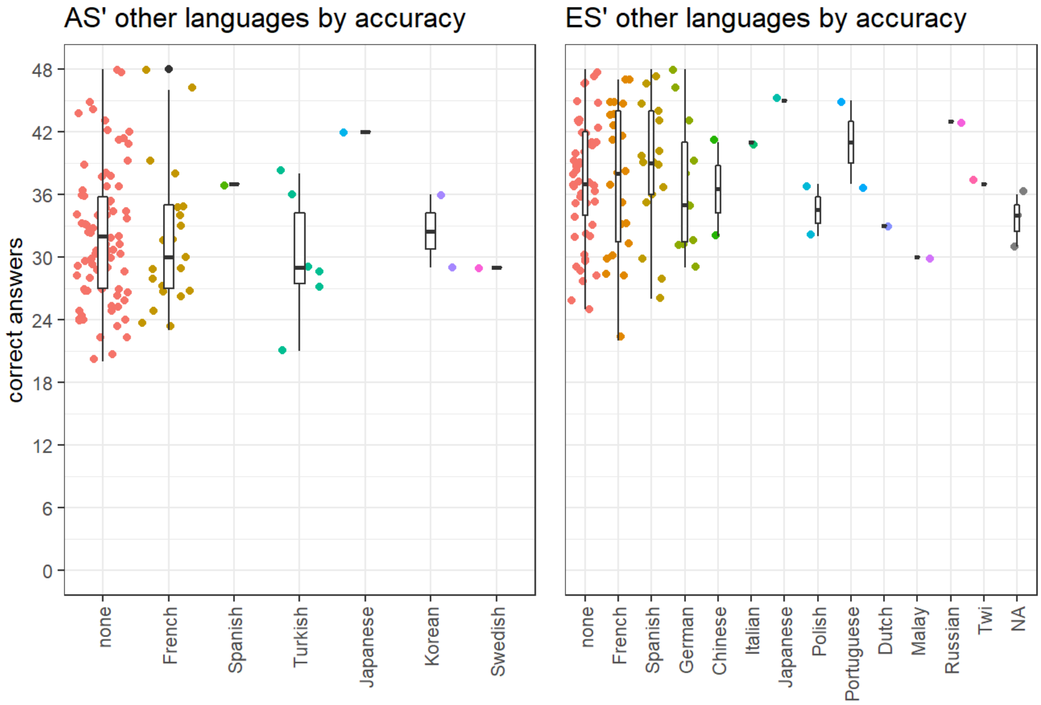


Figure VIII.16: Participant raw accuracy by knowledge of other languages and L1

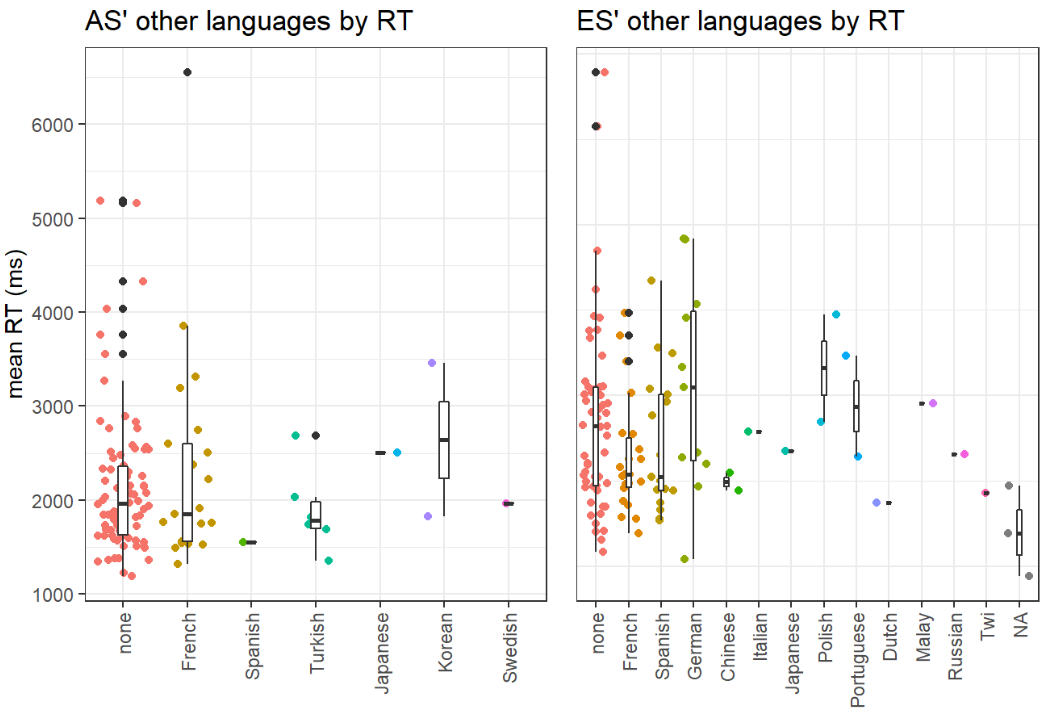


Figure VIII.17: Participant mean RT by knowledge of other languages and L1

To ascertain whether knowledge of certain additional languages (L3) offers an advantage when learning words with different orthographic input (OI), accuracy by L3 and OI condition is plotted. The figure below does not indicate any advantage for the English OI condition for the L1 Arabic group with knowledge of other Roman alphabet languages.

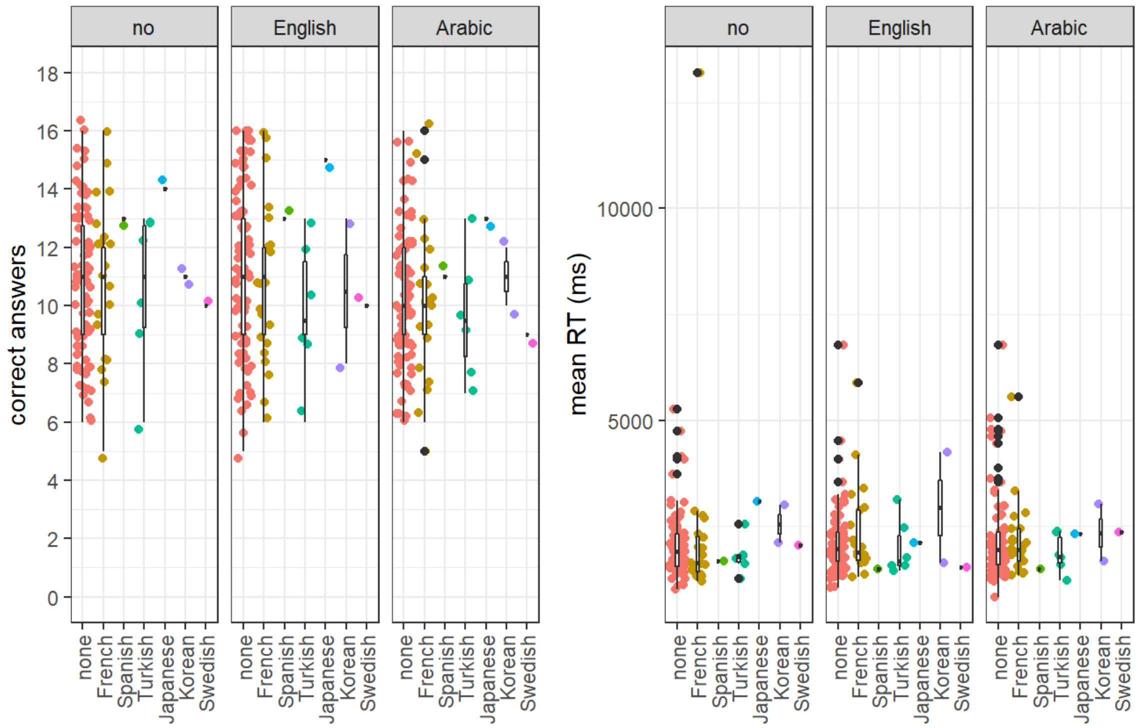


Figure VIII.18: L1 Arabic raw accuracy and mean RT by L3 and OI condition

Arabic-speaking participant L3 experience is then explored in relation to phonological contrasts. L1 Arabic-speakers with experience in French did exhibit marginally higher correct answers, however there is not a notable advantage compared to those who reported no other additional languages.

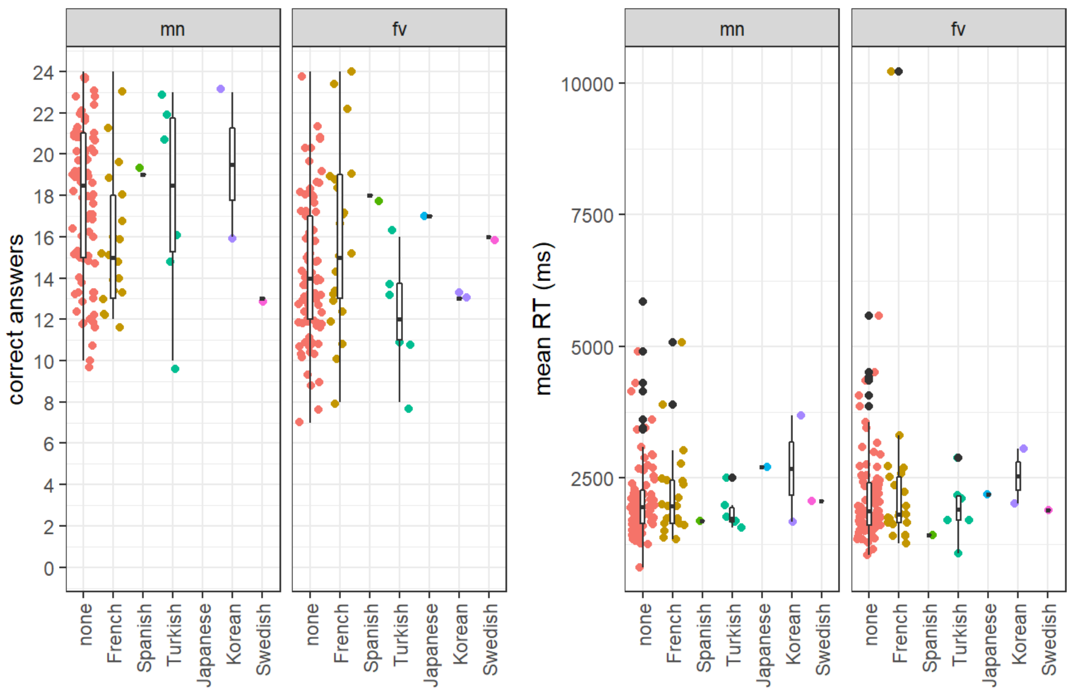


Figure VIII.19: L1 Arabic raw accuracy and mean RT by L3 and phonological contrast

English language exposure and proficiency

This section contains descriptive plots for L1 Arabic-speaking participants related to English language experience and proficiency, including:

1. daily exposure to English
2. age of onset
3. length of study
4. length of stay in Anglophone country
5. self-reported level
6. self-reported skills proficiency (reading, writing, listening, speaking)
7. proficiency test score
8. highest English qualification

The figures below show a positive correlation between raw accuracy and estimated proportion of daily linguistic exposure in English for L1 Arabic-speakers, where increased exposure to English correlates with higher accuracy in the matching task. However, the same effect is not evidenced in response times.

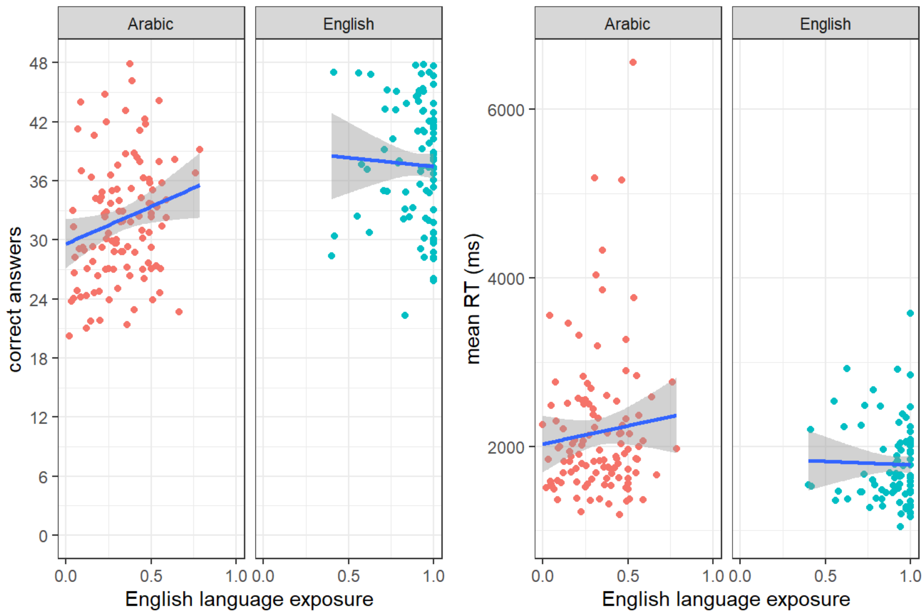


Figure VIII.20: Raw accuracy and mean RT by daily exposure to English

The following plot explores this relationship further in relation to daily exposure to English literacy, which is also positively correlated with accuracy for the L1 Arabic-speakers.

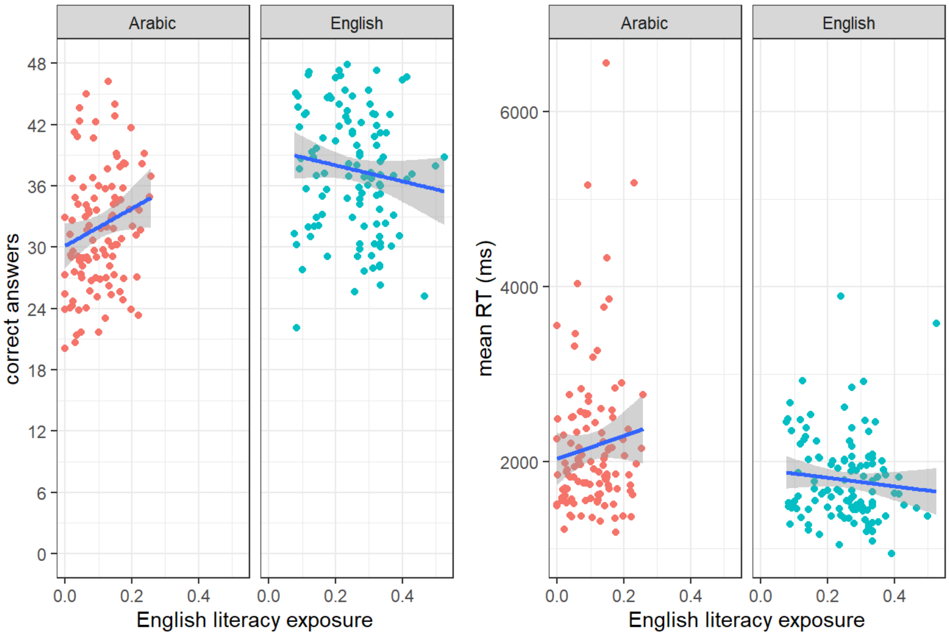


Figure VIII.21: Raw accuracy and mean RT by daily English literacy exposure

To further explore Arabic-speakers' experience and proficiency in English, a new data frame is created with Arabic-speakers only and the English language variables mentioned above. As with the general age of participants, which is likely related to age of onset in learning English, accuracy reduces and reaction times appear to slow as age of onset increases.

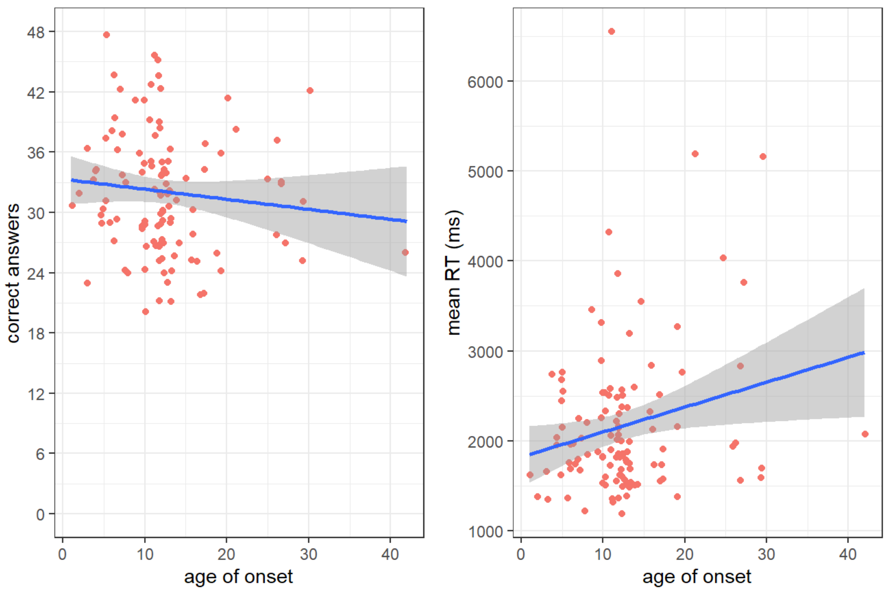


Figure VIII.22: L1 Arabic raw accuracy and mean RT by age of English onset

As would be expected, length of study positively correlates with raw accuracy and negatively correlates with mean RT. As length of study increases, accuracy increases and RT reduces.

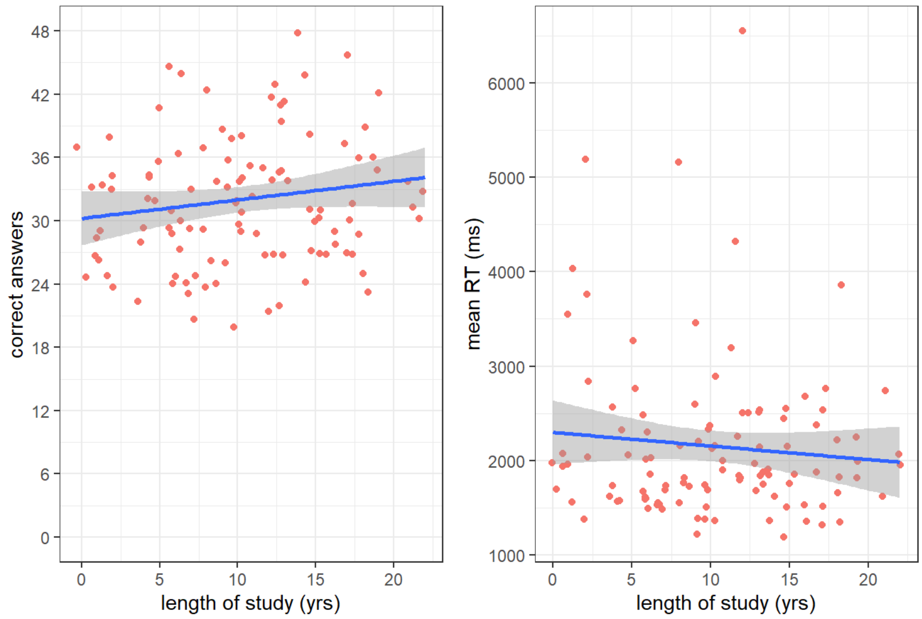


Figure VIII.23: L1 Arabic raw accuracy and mean RT by length of English study

The figures below indicate that the central tendency of accuracy for those who had lived for some time in an Anglophone country is marginally higher than those who had never lived there. This is less evident with reaction times.

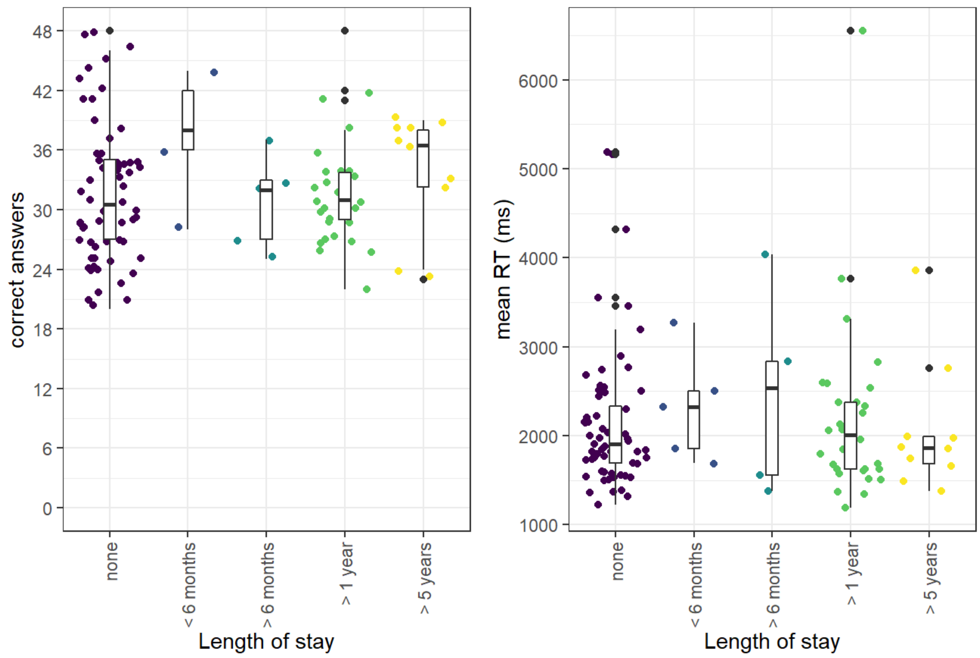


Figure VIII.24: L1 Arabic raw accuracy and mean RT by length of English study

The plots below indicate that accuracy is lower for those who identified as beginner learners of English, with an upward trend in accuracy with higher proficiency; however, this is not reflected in mean RT.

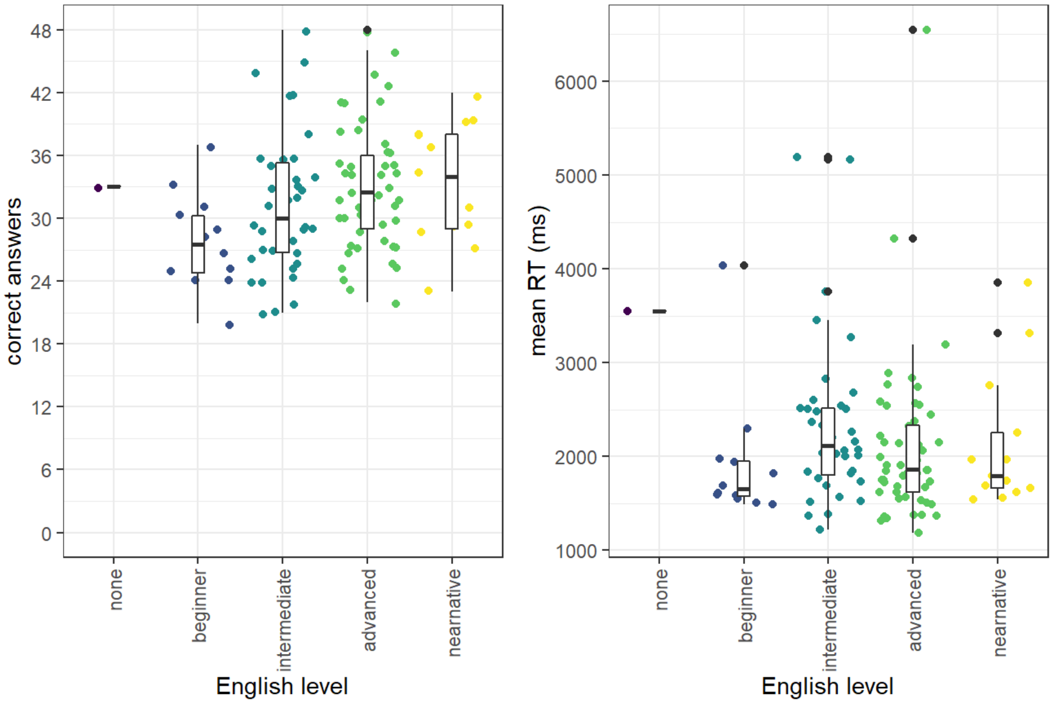


Figure VIII.25: L1 Arabic raw accuracy and mean RT by self-reported level

The figures below plot self-reported ability reading, writing, listening and speaking in English by accuracy and RT. Self-reported English ability appears to have little bearing on RT. Those who reported little to no ability across all skills exhibit lower accuracy in the task, compared to the rest of participants. However, it is also clear from the distribution of participants on the plots that most have intermediate or above ability.

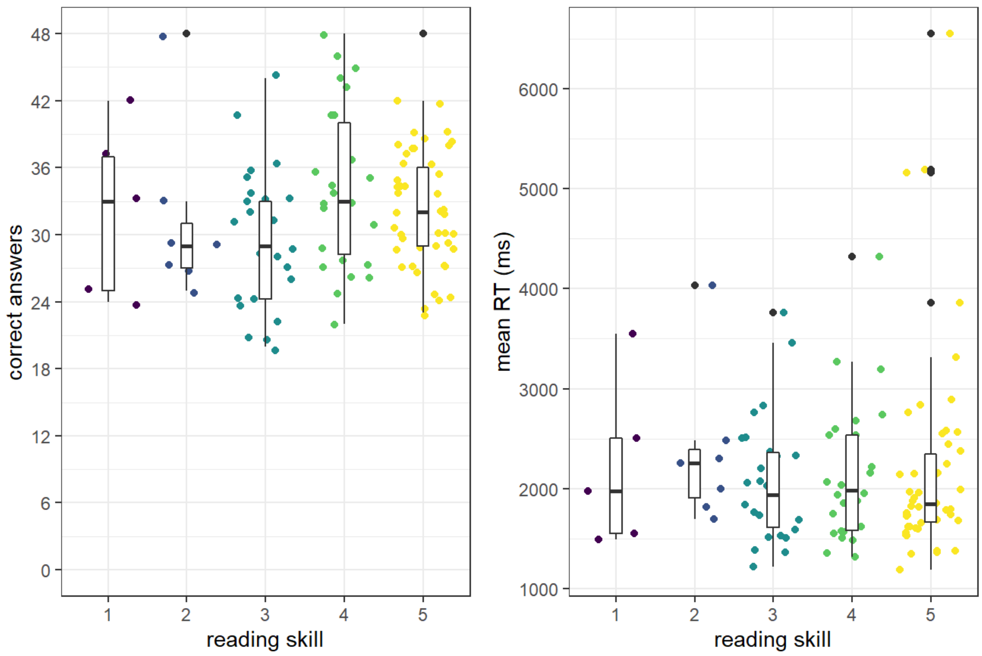


Figure VIII.26: L1 Arabic raw accuracy and mean RT by self-reported reading ability

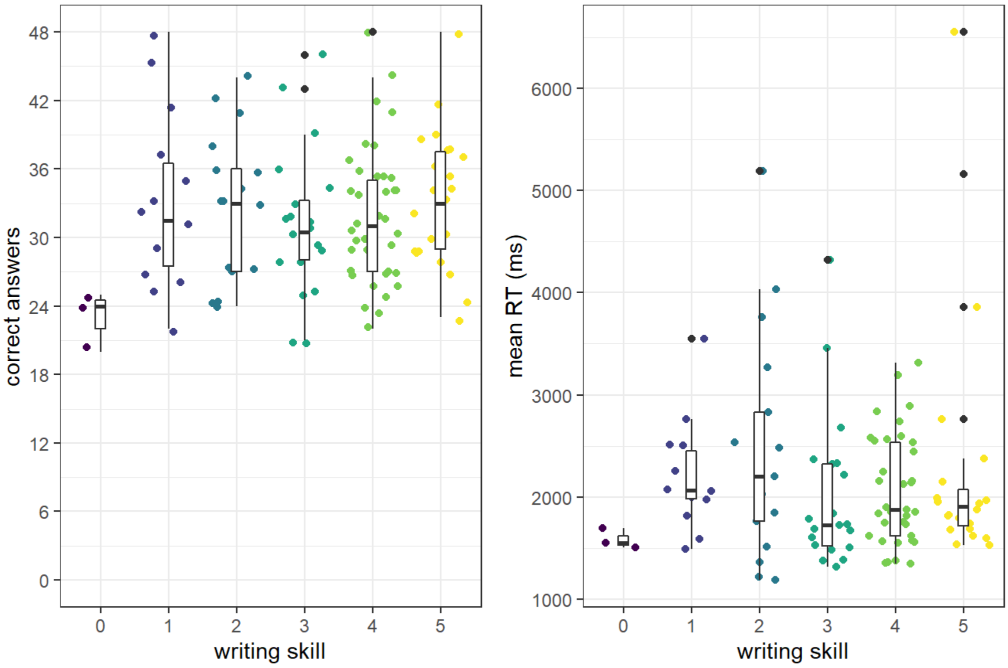


Figure VIII.27: L1 Arabic raw accuracy and mean RT by self-reported writing ability

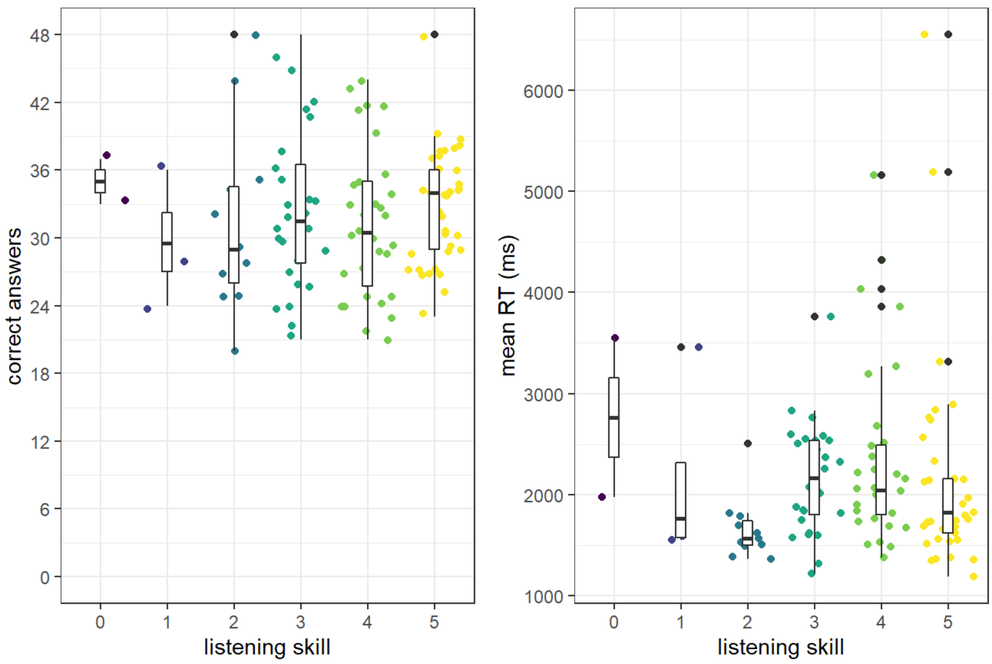


Figure VIII.28: L1 Arabic raw accuracy and mean RT by self-reported listening ability

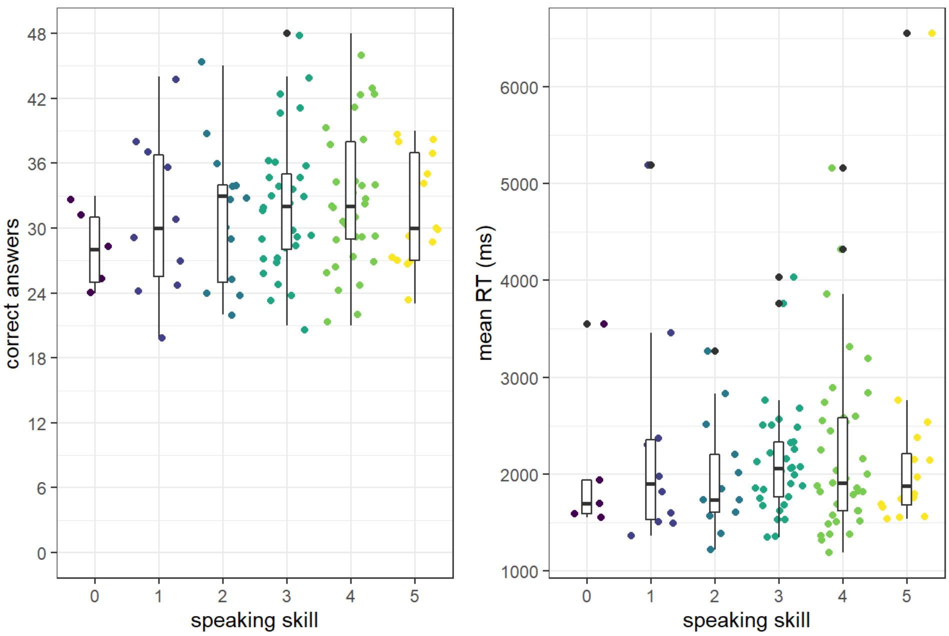


Figure VIII.29: L1 Arabic raw accuracy and mean RT by self-reported speaking ability

Further insight is offered by looking at participants' mean self-reported ability across all four skills. The figures below indicate a slight positive correlation between mean self-reported ability and accuracy, where performance improves for participants with higher ability. However, this is not reflected in reaction times.

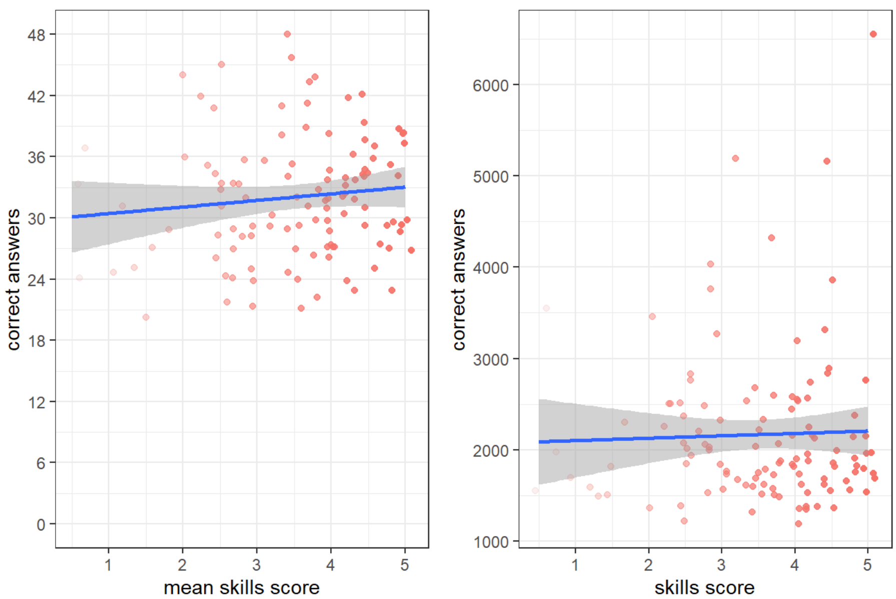


Figure VIII.30: L1 Arabic raw accuracy and mean RT by mean self-reported ability across the four language skills (listening, speaking, reading and writing)

To triangulate self-reported English proficiency measures with a more objective measure of proficiency, L1 Arabic participants completed a short proficiency test. The figures below plot participant scores on the test by accuracy and RT. The test is marked out of 12 and assessed basic vocabulary, grammar, reading and listening in English. Scores on the proficiency test are positively correlated with accuracy and negatively correlated with RT, whereby accuracy increases and RT reduces for participants with higher English proficiency.

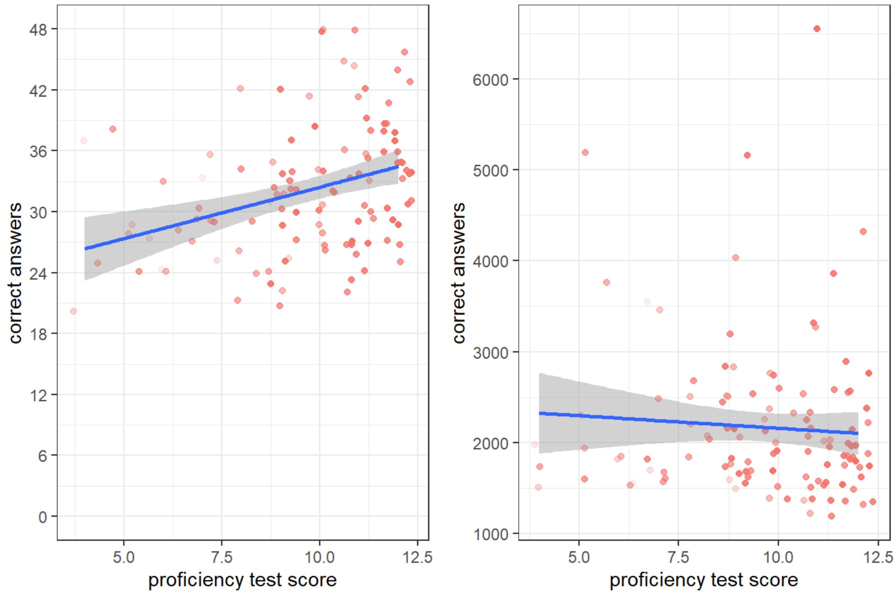


Figure VIII.31: L1 Arabic mean accuracy and RT by English proficiency test score

Finally, the figure below visualises the number of participants who completed different English language qualifications and their performance on the audiovisual matching task. There appears to be a slight advantage for those with English language degrees, which could reflect the proficiency required to complete the qualification and time spent in an Anglophone country. Additionally, it is likely that these participants had an interest in linguistics and had some knowledge of research in this area.

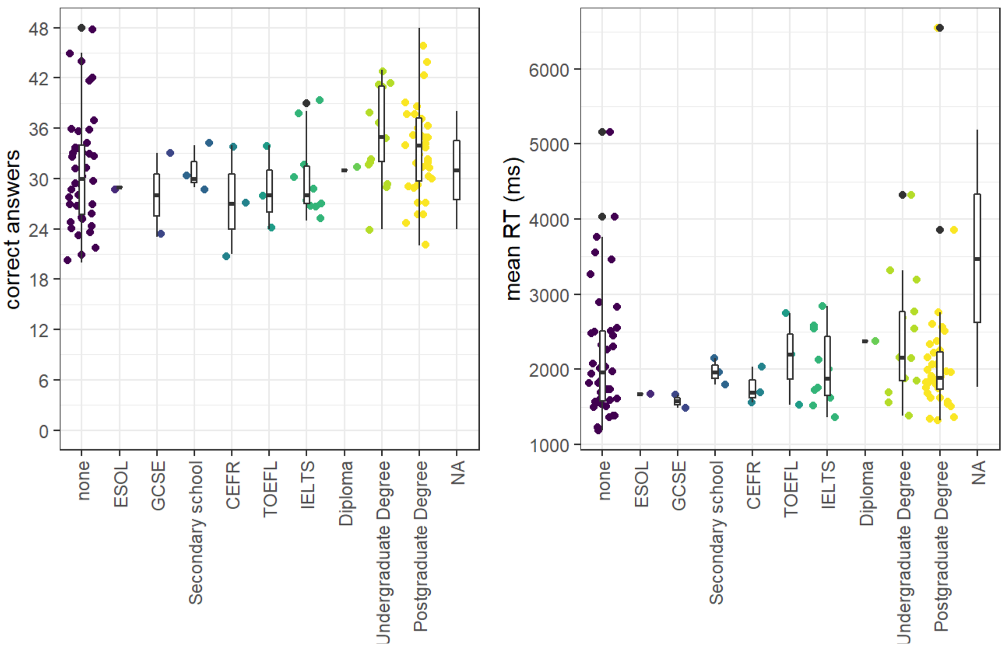


Figure VIII.32: L1 Arabic raw accuracy and mean RT by English language qualification

Appendix IX: Word learning Rmd output

This document outlines the procedure when modelling word learning accuracy and response time using generalised linear mixed-effects models in R (lme4 package).

Accuracy data

Fixed effects of L1, OI and phonological contrast are contrast coded to centre the variables and aid interpretation of the model. The contrast matrices for all factor fixed effects can be found in the output below. The two level variables of language grouping and phonological contrasts are sum coded (L1 Arabic 1, L1 English -1; fv 1, mn -1). Meanwhile, the three level factor of OI is helmert coded to facilitate the comparison between any OI and no OI, then Arabic OI and English OI.

```
#need to relevel the variables
learn_data2$Contrast = relevel(learn_data2$Contrast, ref = "fv")
learn_data2$headphones = relevel(learn_data2$headphones, ref = "no_headphones")
learn_data2$Participant.Device.Type =
relevel(learn_data2$Participant.Device.Type, ref = "computer")

#code categorical predictors
contrasts(learn_data2$Contrast) <- contr.sum(2)
contrasts(learn_data2$Contrast)
##    [,1]
## fv    1
## mn   -1

contrasts(learn_data2$L1) <- contr.sum(2)
contrasts(learn_data2$L1)
##         [,1]
## Arabic     1
## English   -1

contrasts(learn_data$Participant.Device.Type) <- contr.sum(3)
contrasts(learn_data$Participant.Device.Type)
##          [,1] [,2]
## computer    1    0
## mobile      0    1
## tablet     -1   -1

#manually code helmert contrasts as R contr.helmert() codes reverse helmert
contrasts
myhelmert = matrix(c(-2/3, 1/3, 1/3,
                    0, -1/2, 1/2),
                  ncol = 2)

contrasts(learn_data2$OI) = myhelmert
```

```
contrasts(learn_data2$OI)
##                [,1] [,2]
## no      -0.6666667  0.0
## Arabic   0.3333333 -0.5
## English  0.3333333  0.5

contrasts(learn_data2$headphones) = myhelmert
contrasts(learn_data2$headphones)
##                      [,1] [,2]
## no_headphones -0.6666667  0.0
## earphones      0.3333333 -0.5
## headphones     0.3333333  0.5
```

A maximal random effects structure is adopted, which includes random effects with varying intercepts by participant and by item. The maximal model additionally includes random slopes for orthographic input and phonological contrast by participant, however this model does not converge, leading to a model with reduced complexity in the random effects structure. Model comparisons with likelihood tests explore interactions and improved fit with additional fixed effects based on demographic and environmental variables.

```
#gm_learn_01 <- glmer(Correct ~ L1 + OI + Contrast +
 #                  (1 + Contrast + OI | ID) +
  #                 (1 | Item),
   #              data = learn_data2, family = "binomial",
    #             control=glmerControl(optimizer="bobyqa",
optCtrl=list(maxfun=100000)))
#print(summary(gm_learn_01))
# converges but singular fit, reduce complexity of random effects structure

#gm_learn_02 <- glmer(Correct ~ L1 + OI + Contrast +
 #                  (1 + Contrast | ID) +
  #                 (1 | Item),
   #              data = learn_data2, family = "binomial",
    #             control=glmerControl(optimizer="bobyqa",
optCtrl=list(maxfun=100000)))
#print(summary(gm_learn_02))
# does not converge (DNC) with Contrast as a slope in random effects

gm_learn_03 <- glmer(Correct ~ L1 + OI + Contrast +
                   (1 | ID) +
                   (1 | Item),
                 data = learn_data2, family = "binomial",
                 control=glmerControl(optimizer="bobyqa",
optCtrl=list(maxfun=100000)))

print(summary(gm_learn_03))
## Generalized linear mixed model fit by maximum likelihood (Laplace
```

```
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula: Correct ~ L1 + OI + Contrast + (1 | ID) + (1 | Item)
##    Data: learn_data2
## Control: glmerControl(optimizer = "bobyqa", optCtrl = list(maxfun = 1e+05))
##
##      AIC      BIC   logLik deviance df.resid
##   4930.3   4981.6  -2458.2   4916.3    11133
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -7.2263  0.1537  0.1987  0.2662  0.7517
##
## Random effects:
##  Groups Name        Variance Std.Dev.
##  ID     (Intercept) 0.83724  0.9150
##  Item   (Intercept) 0.02693  0.1641
## Number of obs: 11140, groups:  ID, 232; Item, 12
##
## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  3.09444    0.09400  32.921   <2e-16 ***
## L11         -0.23156    0.07527  -3.076   0.0021 **
## OI1         -0.27441    0.13274  -2.067   0.0387 *
## OI2         -0.01115    0.14865  -0.075   0.9402
## Contrast1   -0.02572    0.06167  -0.417   0.6766
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##           (Intr) L11    OI1    OI2
## L11       -0.039
## OI1       -0.032  0.001
## OI2        0.000  0.000 -0.001
## Contrast1 -0.002  0.000 -0.018  0.001
#converges and no singular fit

#anova(gm_learn_01,gm_learn_03)
#significant improvement

#gm_learn_04 <- glmer(Correct ~ L1 * OI + Contrast +
 #                 (1 | ID) +
  #                (1 | Item),
   #              data = learn_data2, family = "binomial",
    #             control=glmerControl(optimizer="bobyqa",
#optCtrl=list(maxfun=100000)))
#print(summary(gm_learn_04))
#converges

#anova(gm_learn_03,gm_learn_04)
```

```
#no improvement

#gm_learn_05 <- glmer(Correct ~ L1 + OI * Contrast +
 #                  (1 | ID) +
  #                 (1 | Item),
   #              data = learn_data2, family = "binomial",
    #             control=glmerControl(optimizer="bobyqa",
optCtrl=list(maxfun=100000)))
#print(summary(gm_learn_05))
#anova(gm_learn_03,gm_learn_05)
#no significant improvement

gm_learn_06 <- glmer(Correct ~ L1 + OI + Contrast + age +
                   (1 | ID) +
                   (1 | Item),
                 data = learn_data2, family = "binomial",
                 control=glmerControl(optimizer="bobyqa",
optCtrl=list(maxfun=100000)))

print(summary(gm_learn_06))
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula: Correct ~ L1 + OI + Contrast + age + (1 | ID) + (1 | Item)
##    Data: learn_data2
## Control: glmerControl(optimizer = "bobyqa", optCtrl = list(maxfun = 1e+05))
##
##      AIC      BIC   logLik deviance df.resid
##   4918.5   4977.1  -2451.3   4902.5    11132
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -7.9512  0.1539  0.2015  0.2684  0.7383
##
## Random effects:
##  Groups Name        Variance Std.Dev.
##  ID     (Intercept) 0.76348  0.8738
##  Item   (Intercept) 0.02685  0.1639
## Number of obs: 11140, groups:  ID, 232; Item, 12
##
## Fixed effects:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  3.981027   0.257869  15.438  < 2e-16 ***
## L11         -0.293571   0.075089  -3.910 9.24e-05 ***
## OI1         -0.274548   0.132753  -2.068 0.038629 *
## OI2         -0.011216   0.148628  -0.075 0.939845
## Contrast1   -0.025703   0.061658  -0.417 0.676779
## age         -0.026634   0.007014  -3.797 0.000146 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Correlation of Fixed Effects:
##           (Intr) L11    OI1    OI2    Cntrs1
## L11       -0.227
## OI1       -0.014  0.002
## OI2        0.000  0.000 -0.001
## Contrast1 -0.001  0.000 -0.018  0.001
## age       -0.934  0.227  0.002  0.000  0.000
#converges and no singular fit

anova(gm_learn_03,gm_learn_06)

## Data: learn_data2
## Models:
## gm_learn_03: Correct ~ L1 + OI + Contrast + (1 | ID) + (1 | Item)
## gm_learn_06: Correct ~ L1 + OI + Contrast + age + (1 | ID) + (1 | Item)
##             npar    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
## gm_learn_03    7 4930.3 4981.6 -2458.2   4916.3
## gm_learn_06    8 4918.5 4977.1 -2451.3   4902.5 13.801  1  0.0002032 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
#significant improvement

## check for environmental and task set up effects

#gm_learn_07 <- glmer(Correct ~ L1 + OI + Contrast + age + prop_distraction +
 #                   (1 | ID) +
  #                  (1 | Item),
   #              data = learn_data2, family = "binomial",
    #             control=glmerControl(optimizer="bobyqa",
optCtrl=list(maxfun=100000)))
#print(summary(gm_learn_07))

#anova(gm_learn_06,gm_learn_07)
#no significant improvement

#gm_learn_08 <- glmer(Correct ~ L1 + OI + Contrast + age + headphones +
 #                   (1 | ID) +
  #                  (1 | Item),
   #              data = learn_data2, family = "binomial",
    #             control=glmerControl(optimizer="bobyqa",
optCtrl=list(maxfun=100000)))
#print(summary(gm_learn_08))

#anova(gm_learn_06, gm_learn_08)
#no significant improvement

#gm_learn_09 <- glmer(Correct ~ L1 + OI + Contrast + age +
Participant.Device.Type +
 #                   (1 | ID) +
```

```
  #                 (1 | Item),
   #              data = learn_data2, family = "binomial",
    #             control=glmerControl(optimizer="bobyqa",
optCtrl=list(maxfun=100000)))
#print(summary(gm_learn_09))

#anova(gm_learn_06, gm_learn_09)
#no significant improvement

## check whether age interacts with L1

#gm_learn_10 <- glmer(Correct ~ OI + Contrast + L1 * age +
 #                  (1 | ID) +
  #                 (1 | Item),
   #              data = learn_data2, family = "binomial",
    #             control=glmerControl(optimizer="bobyqa",
optCtrl=list(maxfun=100000)))
#print(summary(gm_learn_10))

#anova(gm_learn_06, gm_learn_10)
#not significant
```

To assess the contribution of each fixed effect to the fit of the model, further comparisons are performed using likelihood tests between the full final model and depleted versions.

```
#final full model
#gm_learn_06 <- glmer(Correct ~ L1 + OI + Contrast + age +
 #                  (1 | ID) +
  #                 (1 | Item),
   #              data = learn_data2, family = "binomial",
    #             control=glmerControl(optimizer="bobyqa",
optCtrl=list(maxfun=100000)))
#print(summary(gm_learn_06))
#converges and no singular fit

gm_learn_noL1 <- glmer(Correct ~ OI + Contrast + age +
                   (1 | ID) +
                   (1 | Item),
                 data = learn_data2, family = "binomial",
                 control=glmerControl(optimizer="bobyqa",
optCtrl=list(maxfun=100000)))

anova(gm_learn_noL1,gm_learn_06)
## Data: learn_data2
## Models:
## gm_learn_noL1: Correct ~ OI + Contrast + age + (1 | ID) + (1 | Item)
## gm_learn_06: Correct ~ L1 + OI + Contrast + age + (1 | ID) + (1 | Item)
##               npar    AIC    BIC  logLik deviance Chisq Df Pr(>Chisq)
```

```
## gm_learn_noL1    7 4930.9 4982.2 -2458.5   4916.9
## gm_learn_06      8 4918.5 4977.1 -2451.3   4902.5  14.4  1  0.0001478 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
#significant

gm_learn_noOI <- glmer(Correct ~ L1 + Contrast + age +
                   (1 | ID) +
                   (1 | Item),
                 data = learn_data2, family = "binomial",
                 control=glmerControl(optimizer="bobyqa",
optCtrl=list(maxfun=100000)))

anova(gm_learn_noOI,gm_learn_06)
## Data: learn_data2
## Models:
## gm_learn_noOI: Correct ~ L1 + Contrast + age + (1 | ID) + (1 | Item)
## gm_learn_06: Correct ~ L1 + OI + Contrast + age + (1 | ID) + (1 | Item)
##               npar    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
## gm_learn_noOI    6 4918.2 4962.1 -2453.1   4906.2
## gm_learn_06      8 4918.5 4977.1 -2451.3   4902.5 3.6719  2     0.1595
# not significant

gm_learn_noage <- glmer(Correct ~ OI + Contrast + L1 +
                   (1 | ID) +
                   (1 | Item),
                 data = learn_data2, family = "binomial",
                 control=glmerControl(optimizer="bobyqa",
optCtrl=list(maxfun=100000)))

anova(gm_learn_noage,gm_learn_06)
## Data: learn_data2
## Models:
## gm_learn_noage: Correct ~ OI + Contrast + L1 + (1 | ID) + (1 | Item)
## gm_learn_06: Correct ~ L1 + OI + Contrast + age + (1 | ID) + (1 | Item)
##                npar    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
## gm_learn_noage    7 4930.3 4981.6 -2458.2   4916.3
## gm_learn_06       8 4918.5 4977.1 -2451.3   4902.5 13.801  1  0.0002032 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
#significant
```

Response time data

In addition to the fixed effects mentioned in the word learning accuracy model, device type is included in the response time model. This three-level variable is sum coded for comparisons of response times with a computer and mobile phone to the general mean, which includes the third level of tablet response times. As above, L1 and phonological contrast levels are also sum coded. However, in order to fit the model using the Inverse

Gaussian distribution, levels are coded in the opposite direction to above, so that the levels with lower predicted performance are coded as the reference levels (i.e. L1 English 1, L1 Arabic -1; mn 1, fv -1). As before, a maximal random effects structure is adopted.

```
#need to relevel the variables
right_learn_data2$Contrast = relevel(right_learn_data2$Contrast, ref = "mn")
right_learn_data2$L1 = relevel(right_learn_data2$L1, ref = "English")
right_learn_data2$headphones = relevel(right_learn_data2$headphones, ref =
"no_headphones")
right_learn_data2$Participant.Device.Type =
relevel(right_learn_data2$Participant.Device.Type, ref = "computer")

#code categorical predictors
contrasts(right_learn_data2$Contrast) <- contr.sum(2)
contrasts(right_learn_data2$Contrast)
##    [,1]
## mn    1
## fv   -1

contrasts(right_learn_data2$L1) <- contr.sum(2)
contrasts(right_learn_data2$L1)
##         [,1]
## English    1
## Arabic    -1

contrasts(right_learn_data2$Participant.Device.Type) <- contr.sum(3)
contrasts(right_learn_data2$Participant.Device.Type)
##          [,1] [,2]
## computer    1    0
## mobile      0    1
## tablet     -1   -1

#manually code helmert contrasts as R contr.helmert() codes reverse helmert
contrasts
myhelmert = matrix(c(-2/3, 1/3, 1/3,
                     0, -1/2, 1/2),
                   ncol = 2)

contrasts(right_learn_data2$OI) = myhelmert
contrasts(right_learn_data2$OI)
##               [,1] [,2]
## no      -0.6666667  0.0
## Arabic   0.3333333 -0.5
## English  0.3333333  0.5

contrasts(right_learn_data2$headphones) = myhelmert
contrasts(right_learn_data2$headphones)
##                     [,1] [,2]
## no_headphones -0.6666667  0.0
```

```
## earphones      0.3333333 -0.5
## headphones     0.3333333  0.5
```

The model converges with phonological contrast included as a random slope by participants in the random effects structure. As before, model comparisons with likelihood tests explore interactions and improved fit with additional fixed effects based on demographic and environmental variables.

```
#gm_learnRT_01 <- glmer(Reaction.Time ~ L1 + OI + Contrast +
 #                        (1 + Contrast + OI | ID) +
  #                       (1 | Item),
   #                    data = right_learn_data2, family =
inverse.gaussian(link="identity"),
    #                   control = glmerControl(optimizer = "bobyqa", optCtrl =
list(maxfun=100000)))
#print(summary(gm_learnRT_01))
#DNC try with gamma distribution

#gm_learnRT_02 <- glmer(Reaction.Time ~ L1 + OI + Contrast +
 #                        (1 + Contrast + OI | ID) +
  #                       (1 | Item),
   #                    data = right_learn_data2, family = Gamma(link =
"identity"),
    #                   control = glmerControl(optimizer = "bobyqa", optCtrl =
list(maxfun=100000)))
#print(summary(gm_learnRT_02))
#DNC reduce random effects structure back to inverse gaussian

gm_learnRT_03 <- glmer(Reaction.Time ~ L1 + OI + Contrast +
                         (1 + Contrast | ID) +
                         (1 | Item),
                       data = right_learn_data2, family =
inverse.gaussian(link="identity"),
                       control = glmerControl(optimizer = "bobyqa", optCtrl =
list(maxfun=100000)))

print(summary(gm_learnRT_03))
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: inverse.gaussian  ( identity )
## Formula: Reaction.Time ~ L1 + OI + Contrast + (1 + Contrast | ID) + (1 |
##     Item)
##    Data: right_learn_data2
## Control: glmerControl(optimizer = "bobyqa", optCtrl = list(maxfun = 1e+05))
##
##      AIC      BIC   logLik deviance df.resid
## 166574.7 166647.2 -83277.3 166554.7    10437
##
```

```
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -1.8591 -0.5700 -0.2389  0.2451 15.0272
##
## Random effects:
##  Groups   Name        Variance  Std.Dev. Corr
##  ID       (Intercept) 7.750e+04 278.3926
##           Contrast1   1.416e+04 118.9966 -0.06
##  Item     (Intercept) 1.393e+03  37.3257
##  Residual             1.123e-04   0.0106
## Number of obs: 10447, groups:  ID, 232; Item, 12
##
## Fixed effects:
##              Estimate Std. Error t value Pr(>|z|)
## (Intercept)  2190.19      13.46 162.688  < 2e-16 ***
## L11           -46.01      10.37  -4.437 9.11e-06 ***
## OI1            45.82       9.85   4.652 3.30e-06 ***
## OI2           -21.47      11.28  -1.904  0.05693 .
## Contrast1     -32.99      10.98  -3.004  0.00266 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##           (Intr) L11    OI1    OI2
## L11       -0.006
## OI1        0.075  0.095
## OI2        0.214  0.030 -0.063
## Contrast1  0.076  0.018  0.008  0.208
#Converges

#gm_learnRT_04 <- glmer(Reaction.Time ~ L1 + OI + Contrast +
 #                        (1 + OI | ID) +
  #                       (1 | Item),
   #                    data = right_learn_data2, family =
inverse.gaussian(link="identity"),
    #                   control = glmerControl(optimizer = "bobyqa", optCtrl =
list(maxfun=100000)))
#print(summary(gm_learnRT_04))
#singular fit and DNC
```

```
## add interactions

#gm_learnRT_05 <- glmer(Reaction.Time ~ L1 + OI * Contrast +
 #                        (1 + Contrast | ID) +
  #                       (1 | Item),
   #                    data = right_learn_data2, family =
inverse.gaussian(link="identity"),
    #                   control = glmerControl(optimizer = "bobyqa", optCtrl =
list(maxfun=100000)))
```

```
#print(summary(gm_learnRT_05))
#DNC

#gm_learnRT_06 <- glmer(Reaction.Time ~ L1 + OI * Contrast +
 #                          (1 | ID) +
  #                         (1 | Item),
   #                     data = right_learn_data2, family =
inverse.gaussian(link="identity"),
    #                   control = glmerControl(optimizer = "bobyqa", optCtrl =
list(maxfun=100000)))
#print(summary(gm_learnRT_06))
#converges

#anova(gm_learnRT_03,gm_learnRT_07)
#not significant
```

```
## add additional fixed effects

gm_learnRT_07 <- glmer(Reaction.Time ~ L1 + OI + Contrast +
Participant.Device.Type +
                         (1 + Contrast | ID) +
                         (1 | Item),
                       data = right_learn_data2, family =
inverse.gaussian(link="identity"),
                       control = glmerControl(optimizer = "bobyqa", optCtrl =
list(maxfun=100000)))

print(summary(gm_learnRT_07))
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: inverse.gaussian  ( identity )
## Formula: Reaction.Time ~ L1 + OI + Contrast + Participant.Device.Type +
##     (1 + Contrast | ID) + (1 | Item)
##    Data: right_learn_data2
## Control: glmerControl(optimizer = "bobyqa", optCtrl = list(maxfun = 1e+05))
##
##      AIC      BIC   logLik deviance df.resid
## 166573.7 166660.7 -83274.8 166549.7    10435
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -1.8698 -0.5682 -0.2381  0.2430 15.3134
##
## Random effects:
##  Groups   Name        Variance  Std.Dev.  Corr
##  ID       (Intercept) 7.528e+04 274.37809
##           Contrast1   1.411e+04 118.79936 -0.04
##  Item     (Intercept) 1.392e+03  37.30830
##  Residual             1.126e-04   0.01061
```

```
## Number of obs: 10447, groups:  ID, 232; Item, 12
##
## Fixed effects:
##                           Estimate Std. Error t value Pr(>|z|)
## (Intercept)                2075.79      13.82 150.168  < 2e-16 ***
## L11                         -68.74      13.17  -5.219  1.8e-07 ***
## OI1                          45.36      15.65   2.900  0.00374 **
## OI2                         -21.46      11.74  -1.827  0.06771 .
## Contrast1                   -31.45      11.57  -2.717  0.00658 **
## Participant.Device.Type1   177.73      12.92  13.761  < 2e-16 ***
## Participant.Device.Type2    19.06      13.50   1.411  0.15813
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) L11    OI1    OI2    Cntrs1 P.D.T1
## L11         -0.067
## OI1          0.027 -0.193
## OI2         -0.207 -0.034 -0.050
## Contrast1   -0.025 -0.262  0.136 -0.072
## Prtcpn.D.T1  0.009  0.165 -0.151 -0.044 -0.246
## Prtcpn.D.T2  0.152 -0.035  0.011  0.013 -0.138  0.042

anova(gm_learnRT_03,gm_learnRT_07) #marginal significance
## Data: right_learn_data2
## Models:
## gm_learnRT_03: Reaction.Time ~ L1 + OI + Contrast + (1 + Contrast | ID) + (1
| Item)
## gm_learnRT_07: Reaction.Time ~ L1 + OI + Contrast + Participant.Device.Type +
(1 + Contrast | ID) + (1 | Item)
##               npar    AIC    BIC logLik deviance  Chisq Df Pr(>Chisq)
## gm_learnRT_03   10 166575 166647 -83277   166555
## gm_learnRT_07   12 166574 166661 -83275   166550 4.9661  2    0.08349 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

#gm_learnRT_08 <- glmer(Reaction.Time ~ L1 + OI + Contrast + age +
 #                        (1 + Contrast | ID) +
  #                       (1 | Item),
   #                   data = right_learn_data2, family =
inverse.gaussian(link="identity"),
    #                 control = glmerControl(optimizer = "bobyqa", optCtrl =
list(maxfun=100000)))
#print(summary(gm_learnRT_08))
#DNC

#gm_learnRT_09 <- glmer(Reaction.Time ~ L1 + OI + Contrast + prop_distraction +
 #                        (1 + Contrast | ID) +
  #                       (1 | Item),
   #                   data = right_learn_data2, family =
```

```
inverse.gaussian(link="identity"),
    #                   control = glmerControl(optimizer = "bobyqa", optCtrl =
list(maxfun=100000)))
#print(summary(gm_learnRT_9))
#converges

#anova(gm_learnRT_03,gm_learnRT_09)
#not significant

#gm_learnRT_10 <- glmer(Reaction.Time ~ L1 + OI + Contrast + prop_distraction +
Participant.Device.Type +
 #                       (1 + Contrast | ID) +
  #                      (1 | Item),
   #                   data = right_learn_data2, family =
inverse.gaussian(link="identity"),
    #                   control = glmerControl(optimizer = "bobyqa", optCtrl =
list(maxfun=100000)))
#print(summary(gm_learnRT_10))

#anova(gm_learnRT_03,gm_learnRT_10)
#not significant

#gm_learnRT_11 <- glmer(Reaction.Time ~ L1 + OI + Contrast + headphones +
 #                       (1 + Contrast | ID) +
  #                      (1 | Item),
   #                   data = right_learn_data2, family =
inverse.gaussian(link="identity"),
    #                   control = glmerControl(optimizer = "bobyqa", optCtrl =
list(maxfun=100000)))
#print(summary(gm_learnRT_11))

#anova(gm_learnRT_03,gm_learnRT_11)
#not significant
```

To assess the contribution of each fixed effect to the fit of the model, further comparisons are performed using likelihood tests between the full final model and depleted versions.

```
#final model
#gm_learnRT_07 <- glmer(Reaction.Time ~ L1 + OI + Contrast +
#Participant.Device.Type +
 #                       (1 + Contrast | ID) +
  #                      (1 | Item),
   #                   data = right_learn_data2, family =
#inverse.gaussian(link="identity"),
  #                     control = glmerControl(optimizer = "bobyqa", optCtrl =
#list(maxfun=100000)))
#print(summary(gm_learnRT_07))
```

```r
#reduced models
glm_learnRT_noL1 <- glmer(Reaction.Time ~ OI + Contrast +
Participant.Device.Type +
                         (1 + Contrast | ID) +
                         (1 | Item),
                       data = right_learn_data2, family =
inverse.gaussian(link="identity"),
                       control = glmerControl(optimizer = "bobyqa", optCtrl =
list(maxfun=100000)))

anova(glm_learnRT_noL1,gm_learnRT_07)
## Data: right_learn_data2
## Models:
## glm_learnRT_noL1: Reaction.Time ~ OI + Contrast + Participant.Device.Type +
(1 + Contrast | ID) + (1 | Item)
## gm_learnRT_07: Reaction.Time ~ L1 + OI + Contrast + Participant.Device.Type +
(1 + Contrast | ID) + (1 | Item)
##                  npar    AIC    BIC logLik deviance  Chisq Df Pr(>Chisq)
## glm_learnRT_noL1   11 166590 166670 -83284   166568
## gm_learnRT_07      12 166574 166661 -83275   166550 18.349  1  1.839e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
#significant improvement

glm_learnRT_noOI <- glmer(Reaction.Time ~ L1 + Contrast +
Participant.Device.Type +
                         (1 + Contrast | ID) +
                         (1 | Item),
                       data = right_learn_data2, family =
inverse.gaussian(link="identity"),
                       control = glmerControl(optimizer = "bobyqa", optCtrl =
list(maxfun=100000)))

anova(glm_learnRT_noOI,gm_learnRT_07)
## Data: right_learn_data2
## Models:
## glm_learnRT_noOI: Reaction.Time ~ L1 + Contrast + Participant.Device.Type +
(1 + Contrast | ID) + (1 | Item)
## gm_learnRT_07: Reaction.Time ~ L1 + OI + Contrast + Participant.Device.Type +
(1 + Contrast | ID) + (1 | Item)
##                  npar    AIC    BIC logLik deviance  Chisq Df Pr(>Chisq)
## glm_learnRT_noOI   10 166571 166644 -83276   166551
## gm_learnRT_07      12 166574 166661 -83275   166550 1.3774  2     0.5022
# no significant improvement

glm_learnRT_noContrast <- glmer(Reaction.Time ~ L1 + OI +
Participant.Device.Type +
                         (1 + Contrast | ID) +
                         (1 | Item),
                       data = right_learn_data2, family =
```

```
inverse.gaussian(link="identity"),
                       control = glmerControl(optimizer = "bobyqa", optCtrl =
list(maxfun=100000)))

anova(glm_learnRT_noContrast,gm_learnRT_07)
## Data: right_learn_data2
## Models:
## glm_learnRT_noContrast: Reaction.Time ~ L1 + OI + Participant.Device.Type +
(1 + Contrast | ID) + (1 | Item)
## gm_learnRT_07: Reaction.Time ~ L1 + OI + Contrast + Participant.Device.Type +
(1 + Contrast | ID) + (1 | Item)
##                        npar    AIC    BIC logLik deviance  Chisq Df
Pr(>Chisq)
## glm_learnRT_noContrast   11 166588 166668 -83283   166566
## gm_learnRT_07            12 166574 166661 -83275   166550 16.013  1
6.29e-05
##
## glm_learnRT_noContrast
## gm_learnRT_07          ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
# significant improvement

glm_learnRT_nodevice <- glmer(Reaction.Time ~ L1 + OI + Contrast +
                         (1 + Contrast | ID) +
                         (1 | Item),
                       data = right_learn_data2, family =
inverse.gaussian(link="identity"),
                       control = glmerControl(optimizer = "bobyqa", optCtrl =
list(maxfun=100000)))

anova(glm_learnRT_nodevice,gm_learnRT_07)
## Data: right_learn_data2
## Models:
## glm_learnRT_nodevice: Reaction.Time ~ L1 + OI + Contrast + (1 + Contrast |
ID) + (1 | Item)
## gm_learnRT_07: Reaction.Time ~ L1 + OI + Contrast + Participant.Device.Type +
(1 + Contrast | ID) + (1 | Item)
##                      npar    AIC    BIC logLik deviance  Chisq Df Pr(>Chisq)
## glm_learnRT_nodevice   10 166575 166647 -83277   166555
## gm_learnRT_07          12 166574 166661 -83275   166550 4.9661  2    0.08349
.
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
#approaching significance
```

Appendix X: Word matching GLMM Rmd output

This document outlines the procedure when modelling accuracy data from the audio-visual matching task using generalised linear mixed-effects models in R (lme4 package).

L1 English accuracy data

Fixed effects of trial type, OI and phonological contrast are contrast coded to centre the variables and aid interpretation of the model. The two-level variables of trial type and phonological contrasts are sum coded (mismatch 1, match -1; fv 1, mn -1). Meanwhile, the three-level factor of OI is helmert coded to facilitate the comparison between any OI and no OI, then Arabic OI and English OI.

```
#need to relevel the variables
ES_data$Contrast = relevel(ES_data$Contrast, ref = "fv")
ES_data$Match = relevel(ES_data$Match, ref = "mismatch")
ES_data$OI = relevel(ES_data$OI, ref = "no")
ES_data$headphones = relevel(ES_data$headphones, ref = "no_headphones")
ES_data$Participant.Device.Type = relevel(ES_data$Participant.Device.Type, ref =
"computer")

#code categorical predictors
contrasts(ES_data$Match) <- contr.sum(2)
contrasts(ES_data$Match)
##          [,1]
## mismatch    1
## match      -1

contrasts(ES_data$Contrast) <- contr.sum(2)
contrasts(ES_data$Contrast)
##    [,1]
## fv    1
## mn   -1

contrasts(ES_data$Participant.Device.Type) <- contr.sum(3)
contrasts(ES_data$Participant.Device.Type)
##          [,1] [,2]
## computer    1    0
## mobile      0    1
## tablet     -1   -1

#polynomial contrasts for education
contrasts (ES_data$education) <- contr.poly(6)
contrasts (ES_data$education)
##                    .L         .Q         .C         ^4          ^5
## primary    -0.5976143  0.5455447 -0.3726780  0.1889822 -0.06299408
## secondary  -0.3585686 -0.1091089  0.5217492 -0.5669467  0.31497039
## other_prof -0.1195229 -0.4364358  0.2981424  0.3779645 -0.62994079
```

```
## bachelors   0.1195229 -0.4364358 -0.2981424  0.3779645  0.62994079
## masters     0.3585686 -0.1091089 -0.5217492 -0.5669467 -0.31497039
## doctorate   0.5976143  0.5455447  0.3726780  0.1889822  0.06299408

#manually code helmert contrasts as R contr.helmert() codes reverse helmert
contrasts
myhelmert = matrix(c(-2/3, 1/3, 1/3,
                     0, -1/2, 1/2),
                   ncol = 2)

contrasts(ES_data$OI) = myhelmert
contrasts(ES_data$OI)
##               [,1] [,2]
## no      -0.6666667  0.0
## Arabic   0.3333333 -0.5
## English  0.3333333  0.5

contrasts(ES_data$headphones) = myhelmert
contrasts(ES_data$headphones)
##                     [,1] [,2]
## no_headphones -0.6666667  0.0
## earphones      0.3333333 -0.5
## headphones     0.3333333  0.5
```

As with the word learning analysis, a stepwise procedure is followed to build the model, whereby the theoretically motivated fixed effects of trial type (match vs. mismatch), OI and phonological contrast are initially added to the model. A maximal random effects structure is adopted, including random effects with varying intercepts by participant and by item. The maximal model also includes random slopes for trial type, orthographic input and phonological contrast by participant, as well as trial type by item. However, this model does not converge, leading to a model with reduced complexity in the random effects structure. Model comparisons with likelihood tests explore interactions and improved fit with additional fixed effects based on demographic and environmental variables.

```
#base model
#gm_en01 <- glmer(Correct ~ OI + Contrast + Match +
 #                  (1 + Contrast + OI + Match| ID) +
  #                 (1 + Match | Item),
   #              data = ES_data, family = "binomial",
    #             control=glmerControl(optimizer="bobyqa",
optCtrl=list(maxfun=100000)))
#print(summary(gm_en01))
#singular fit - reduce random effects structure

#gm_en02 <- glmer(Correct ~ OI + Contrast + Match +
 #                  (1 + Contrast + OI| ID) +
```

```
  #                 (1 + Match |Item),
   #              data = ES_data, family = "binomial",
    #             control=glmerControl(optimizer="bobyqa",
optCtrl=list(maxfun=100000)))
#print(summary(gm_en02))
#singular fit - reduce random effects structure

gm_en03 <- glmer(Correct ~ OI + Contrast + Match +
                   (1 + Contrast | ID) +
                   (1 + Match | Item),
                 data = ES_data, family = "binomial",
                 control=glmerControl(optimizer="bobyqa",
optCtrl=list(maxfun=100000)))

print(summary(gm_en03))
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula: Correct ~ OI + Contrast + Match + (1 + Contrast | ID) + (1 +
##     Match | Item)
##    Data: ES_data
## Control: glmerControl(optimizer = "bobyqa", optCtrl = list(maxfun = 1e+05))
##
##      AIC      BIC   logLik deviance df.resid
##   5011.2   5084.3  -2494.6   4989.2     5653
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -6.8691  0.1253  0.3000  0.4833  1.7971
##
## Random effects:
##  Groups Name        Variance Std.Dev. Corr
##  ID     (Intercept) 0.97265  0.9862
##         Contrast1   0.16141  0.4018   -0.37
##  Item   (Intercept) 0.04943  0.2223
##         Match1      0.05900  0.2429   -0.08
## Number of obs: 5664, groups:  ID, 118; Item, 12
##
## Fixed effects:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  1.77597    0.12151  14.615  < 2e-16 ***
## OI1         -0.05713    0.17274  -0.331    0.741
## OI2          0.25833    0.18245   1.416    0.157
## Contrast1   -0.46113    0.09272  -4.974 6.57e-07 ***
## Match1      -0.79667    0.08085  -9.854  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##           (Intr) OI1    OI2    Cntrs1
```

```
## OI1         0.006
## OI2         0.015  0.041
## Contrast1 -0.187 -0.148 -0.037
## Match1    -0.112 -0.013 -0.014  0.038
# converges and no singular fit

gm_en_reduced <- glmer(Correct ~ 1 +
                         (1 + Contrast | ID) +
                         (1 + Match | Item),
                       data = ES_data, family = "binomial",
                       control=glmerControl(optimizer="bobyqa",
optCtrl=list(maxfun=100000)))

anova(gm_en_reduced,gm_en03)
## Data: ES_data
## Models:
## gm_en_reduced: Correct ~ 1 + (1 + Contrast | ID) + (1 + Match | Item)
## gm_en03: Correct ~ OI + Contrast + Match + (1 + Contrast | ID) + (1 + Match |
Item)
##               npar    AIC    BIC  logLik deviance Chisq Df Pr(>Chisq)
## gm_en_reduced    7 5045.0 5091.5 -2515.5   5031.0
## gm_en03         11 5011.2 5084.3 -2494.6   4989.2 41.74  4  1.889e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
#significant improvement on reduced model
## adding in interactions
gm_en04 <- glmer(Correct ~ OI * Contrast + Match +
                   (1 + Contrast | ID) +
                   (1 + Match | Item),
                 data = ES_data, family = "binomial",
                 control=glmerControl(optimizer="bobyqa",
optCtrl=list(maxfun=100000)))
#print(summary(gm_en04))

anova(gm_en03, gm_en04)
## Data: ES_data
## Models:
## gm_en03: Correct ~ OI + Contrast + Match + (1 + Contrast | ID) + (1 + Match |
Item)
## gm_en04: Correct ~ OI * Contrast + Match + (1 + Contrast | ID) + (1 + Match |
Item)
##         npar    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
## gm_en03   11 5011.2 5084.3 -2494.6   4989.2
## gm_en04   13 5006.1 5092.4 -2490.1   4980.1 9.1434  2    0.01034 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
#significant improvement

#gm_en05 <- glmer(Correct ~ OI * Contrast * Match +
 #                  (1 + Contrast | ID) +
```

```
  #                 (1 + Match | Item),
   #              data = ES_data, family = "binomial",
    #             control=glmerControl(optimizer="bobyqa",
optCtrl=list(maxfun=100000)))
#print(summary(gm_en05))
#anova(gm_en04, gm_en05)
#not a significant improvement

##add additional fixed effects
#gm_en06 <- glmer(Correct ~ OI * Contrast + Match + prop_distraction +
 #                  (1 + Contrast | ID) +
  #                 (1 + Match | Item),
   #              data = ES_data, family = "binomial",
    #             control=glmerControl(optimizer="bobyqa",
optCtrl=list(maxfun=100000)))
#print(summary(gm_en06))

#anova(gm_en04, gm_en06)
#marginal improvement

gm_en07 <- glmer(Correct ~ OI * Contrast + Match + age +
                   (1 + Contrast | ID) +
                   (1 + Match | Item),
                 data = ES_data, family = "binomial",
                 control=glmerControl(optimizer="bobyqa",
optCtrl=list(maxfun=100000)))

print(summary(gm_en07))
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula: Correct ~ OI * Contrast + Match + age + (1 + Contrast | ID) +
##     (1 + Match | Item)
##    Data: ES_data
## Control: glmerControl(optimizer = "bobyqa", optCtrl = list(maxfun = 1e+05))
##
##      AIC      BIC   logLik deviance df.resid
##   4998.8   5091.8  -2485.4   4970.8     5650
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -6.4263  0.1241  0.2989  0.4848  1.7819
##
## Random effects:
##  Groups Name        Variance Std.Dev. Corr
##  ID     (Intercept) 0.87054  0.9330
##         Contrast1   0.15919  0.3990   -0.35
##  Item   (Intercept) 0.01716  0.1310
##         Match1      0.06058  0.2461   -0.35
## Number of obs: 5664, groups:  ID, 118; Item, 12
```

```
##
## Fixed effects:
##                Estimate Std. Error z value Pr(>|z|)
## (Intercept)    2.580606   0.282672   9.129  < 2e-16 ***
## OI1           -0.026358   0.118875  -0.222 0.824523
## OI2            0.269011   0.126540   2.126 0.033512 *
## Contrast1     -0.465613   0.071932  -6.473 9.61e-11 ***
## Match1        -0.798557   0.081698  -9.774  < 2e-16 ***
## age           -0.022898   0.007374  -3.105 0.001902 **
## OI1:Contrast1 -0.395582   0.109882  -3.600 0.000318 ***
## OI2:Contrast1 -0.111692   0.131843  -0.847 0.396908
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) OI1    OI2    Cntrs1 Match1 age    OI1:C1
## OI1          0.000
## OI2          0.012  0.036
## Contrast1   -0.071 -0.158 -0.037
## Match1      -0.085 -0.009 -0.016  0.038
## age         -0.927  0.004 -0.003 -0.025  0.013
## OI1:Cntrst1 -0.017 -0.098 -0.027  0.040  0.021  0.004
## OI2:Cntrst1 -0.004  0.092 -0.107 -0.057 -0.003 -0.001 -0.006

anova(gm_en04, gm_en07)
## Data: ES_data
## Models:
## gm_en04: Correct ~ OI * Contrast + Match + (1 + Contrast | ID) + (1 + Match | 
Item)
## gm_en07: Correct ~ OI * Contrast + Match + age + (1 + Contrast | ID) + (1 + 
Match | Item)
##         npar    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
## gm_en04   13 5006.1 5092.4 -2490.1   4980.1
## gm_en07   14 4998.8 5091.8 -2485.4   4970.8 9.3279  1   0.002257 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
#significant improvement

#gm_en08 <- glmer(Correct ~ OI * Contrast + Match + age + headphones +
 #                 (1 + Contrast | ID) +
  #                (1 + Match | Item),
   #             data = ES_data, family = "binomial",
    #            control=glmerControl(optimizer="bobyqa",
optCtrl=list(maxfun=100000)))
#print(summary(gm_en08))

#anova(gm_en07, gm_en08)
#no significant improvement

#gm_en09 <- glmer(Correct ~ OI * Contrast + Match + age +
```

```
Participant.Device.Type +
 #                  (1 + Contrast | ID) +
  #                 (1 + Match | Item),
   #              data = ES_data, family = "binomial",
    #             control=glmerControl(optimizer="bobyqa",
optCtrl=list(maxfun=100000)))
#print(summary(gm_en09))

#anova(gm_en07, gm_en09)
#no significant improvement

#gm_en10 <- glmer(Correct ~ OI * Contrast + Match + education +
 #                  (1 + Contrast | ID) +
  #                 (1 + Match | Item),
   #              data = ES_data, family = "binomial",
    #             control=glmerControl(optimizer="bobyqa",
optCtrl=list(maxfun=100000)))
#print(summary(gm_en10))

#anova(gm_en07, gm_en10)
#no significant improvement
```

To assess the contribution of each fixed effect to the fit of the model, further comparisons are performed using likelihood tests between the full final model and depleted versions.

```
#final full model
#gm_en07 <- glmer(Correct ~ OI * Contrast + Match + age +
 #                  (1 + Contrast | ID) +
  #                 (1 + Match | Item),
   #              data = ES_data, family = "binomial",
    #             control=glmerControl(optimizer="bobyqa",
optCtrl=list(maxfun=100000)))
#print(summary(gm_en07))

gm_enmatch_noOI <- glmer(Correct ~ Contrast + Match + age +
                   (1 + Contrast | ID) +
                   (1 + Match | Item),
                 data = ES_data, family = "binomial",
                 control=glmerControl(optimizer="bobyqa",
optCtrl=list(maxfun=100000)))

anova(gm_enmatch_noOI,gm_en07)
## Data: ES_data
## Models:
## gm_enmatch_noOI: Correct ~ Contrast + Match + age + (1 + Contrast | ID) + (1
+ Match | Item)
## gm_en07: Correct ~ OI * Contrast + Match + age + (1 + Contrast | ID) + (1 +
Match | Item)
```

```
##                  npar    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
## gm_enmatch_noOI   10 5001.9 5068.3 -2491.0   4981.9
## gm_en07           14 4998.8 5091.8 -2485.4   4970.8 11.151  4    0.02492 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
#significant

gm_enmatch_noContrast <- glmer(Correct ~ OI + Match + age +
                   (1 + Contrast | ID) +
                   (1 + Match | Item),
                 data = ES_data, family = "binomial",
                 control=glmerControl(optimizer="bobyqa",
optCtrl=list(maxfun=100000)))

anova(gm_enmatch_noContrast,gm_en07)
## Data: ES_data
## Models:
## gm_enmatch_noContrast: Correct ~ OI + Match + age + (1 + Contrast | ID) + (1
+ Match | Item)
## gm_en07: Correct ~ OI * Contrast + Match + age + (1 + Contrast | ID) + (1 +
Match | Item)
##                        npar    AIC    BIC  logLik deviance Chisq Df Pr(>Chisq)
## gm_enmatch_noContrast   11 5016.3 5089.3 -2497.1   4994.3
## gm_en07                 14 4998.8 5091.8 -2485.4   4970.8 23.51  3  3.161e-05
##
## gm_enmatch_noContrast
## gm_en07                ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
# highly significant

gm_enmatch_noMatch <- glmer(Correct ~ OI * Contrast + age +
                   (1 + Contrast | ID) +
                   (1 + Match | Item),
                 data = ES_data, family = "binomial",
                 control=glmerControl(optimizer="bobyqa",
optCtrl=list(maxfun=100000)))

anova(gm_enmatch_noMatch,gm_en07)
## Data: ES_data
## Models:
## gm_enmatch_noMatch: Correct ~ OI * Contrast + age + (1 + Contrast | ID) + (1
+ Match | Item)
## gm_en07: Correct ~ OI * Contrast + Match + age + (1 + Contrast | ID) + (1 +
Match | Item)
##                     npar    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
## gm_enmatch_noMatch    13 5022.9 5109.3 -2498.5   4996.9
## gm_en07               14 4998.8 5091.8 -2485.4   4970.8 26.145  1  3.167e-07
***
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
# highly significant

gm_enmatch_noage <- glmer(Correct ~ OI * Contrast + Match +
                   (1 + Contrast | ID) +
                   (1 + Match | Item),
                 data = ES_data, family = "binomial",
                 control=glmerControl(optimizer="bobyqa",
optCtrl=list(maxfun=100000)))

anova(gm_enmatch_noage,gm_en07)
## Data: ES_data
## Models:
## gm_enmatch_noage: Correct ~ OI * Contrast + Match + (1 + Contrast | ID) + (1
+ Match | Item)
## gm_en07: Correct ~ OI * Contrast + Match + age + (1 + Contrast | ID) + (1 +
Match | Item)
##                  npar    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
## gm_enmatch_noage   13 5006.1 5092.4 -2490.1   4980.1
## gm_en07            14 4998.8 5091.8 -2485.4   4970.8 9.3279  1   0.002257 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
#significant

gm_enmatch_noInteract <- glmer(Correct ~ OI + Contrast + Match + age +
                   (1 + Contrast | ID) +
                   (1 + Match | Item),
                 data = ES_data, family = "binomial",
                 control=glmerControl(optimizer="bobyqa",
optCtrl=list(maxfun=100000)))

anova(gm_enmatch_noInteract,gm_en07)
## Data: ES_data
## Models:
## gm_enmatch_noInteract: Correct ~ OI + Contrast + Match + age + (1 + Contrast
| ID) + (1 + Match | Item)
## gm_en07: Correct ~ OI * Contrast + Match + age + (1 + Contrast | ID) + (1 +
Match | Item)
##                       npar    AIC    BIC  logLik deviance  Chisq Df
Pr(>Chisq)
## gm_enmatch_noInteract   12 5003.9 5083.6 -2490.0   4979.9
## gm_en07                 14 4998.8 5091.8 -2485.4   4970.8 9.1538  2
0.01029
##
## gm_enmatch_noInteract
## gm_en07               *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
# significant
```

Fixed effects of trial type, OI and phonological contrast are contrast coded to centre the variables and aid interpretation of the model. The two-level variables of trial type and phonological contrasts are sum coded (mismatch 1, match -1; fv 1, mn -1). Meanwhile, the three-level factor of OI is helmert coded to facilitate the comparison between any OI and no OI, then Arabic OI and English OI.

```
#need to relevel the variables
ES_data$Contrast = relevel(ES_data$Contrast, ref = "fv")
ES_data$Match = relevel(ES_data$Match, ref = "mismatch")
ES_data$OI = relevel(ES_data$OI, ref = "no")
ES_data$headphones = relevel(ES_data$headphones, ref = "no_headphones")
ES_data$Participant.Device.Type = relevel(ES_data$Participant.Device.Type, ref =
"computer")

#code categorical predictors
contrasts(ES_data$Match) <- contr.sum(2)
contrasts(ES_data$Match)
##          [,1]
## mismatch    1
## match      -1

contrasts(ES_data$Contrast) <- contr.sum(2)
contrasts(ES_data$Contrast)
##    [,1]
## fv    1
## mn   -1

contrasts(ES_data$Participant.Device.Type) <- contr.sum(3)
contrasts(ES_data$Participant.Device.Type)
##          [,1] [,2]
## computer    1    0
## mobile      0    1
## tablet     -1   -1

#polynomial contrasts for education
contrasts (ES_data$education) <- contr.poly(6)
contrasts (ES_data$education)
##                    .L         .Q         .C         ^4          ^5
## primary    -0.5976143  0.5455447 -0.3726780  0.1889822 -0.06299408
## secondary  -0.3585686 -0.1091089  0.5217492 -0.5669467  0.31497039
## other_prof -0.1195229 -0.4364358  0.2981424  0.3779645 -0.62994079
## bachelors   0.1195229 -0.4364358 -0.2981424  0.3779645  0.62994079
## masters     0.3585686 -0.1091089 -0.5217492 -0.5669467 -0.31497039
## doctorate   0.5976143  0.5455447  0.3726780  0.1889822  0.06299408

#manually code helmert contrasts as R contr.helmert() codes reverse helmert
```

```
contrasts
myhelmert = matrix(c(-2/3, 1/3, 1/3,
                     0, -1/2, 1/2),
                   ncol = 2)

contrasts(ES_data$OI) = myhelmert
contrasts(ES_data$OI)
##               [,1] [,2]
## no      -0.6666667  0.0
## Arabic   0.3333333 -0.5
## English  0.3333333  0.5

contrasts(ES_data$headphones) = myhelmert
contrasts(ES_data$headphones)
##                     [,1] [,2]
## no_headphones -0.6666667  0.0
## earphones      0.3333333 -0.5
## headphones     0.3333333  0.5
```

As before, a stepwise procedure is followed to build the model, whereby the theoretically motivated fixed effects of trial type (match vs. mismatch), OI and phonological contrast are initially added to the model. A maximal random effects structure is adopted, including random effects with varying intercepts by participant and by item. The maximal model also includes random slopes for trial type, OI and phonological contrast by participant, as well as trial type by item. However, this model does not converge, leading to a model with reduced complexity in the random effects structure. Model comparisons with likelihood tests explore interactions and improved fit with additional fixed effects based on demographic and environmental variables.

```
#base model
#gm_en01 <- glmer(Correct ~ OI + Contrast + Match +
 #                  (1 + Contrast + OI + Match| ID) +
  #                 (1 + Match | Item),
   #             data = ES_data, family = "binomial",
    #            control=glmerControl(optimizer="bobyqa",
optCtrl=list(maxfun=100000)))
#print(summary(gm_en01))
#singular fit - reduce random effects structure

#gm_en02 <- glmer(Correct ~ OI + Contrast + Match +
 #                  (1 + Contrast + OI| ID) +
  #                 (1 + Match |Item),
   #             data = ES_data, family = "binomial",
    #            control=glmerControl(optimizer="bobyqa",
optCtrl=list(maxfun=100000)))
#print(summary(gm_en02))
```

```
#singular fit - reduce random effects structure

gm_en03 <- glmer(Correct ~ OI + Contrast + Match +
                   (1 + Contrast | ID) +
                   (1 + Match | Item),
                 data = ES_data, family = "binomial",
                 control=glmerControl(optimizer="bobyqa",
optCtrl=list(maxfun=100000)))

print(summary(gm_en03))
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula: Correct ~ OI + Contrast + Match + (1 + Contrast | ID) + (1 +
##     Match | Item)
##    Data: ES_data
## Control: glmerControl(optimizer = "bobyqa", optCtrl = list(maxfun = 1e+05))
##
##      AIC      BIC   logLik deviance df.resid
##   5011.2   5084.3  -2494.6   4989.2     5653
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -6.8691  0.1253  0.3000  0.4833  1.7971
##
## Random effects:
##  Groups Name        Variance Std.Dev. Corr
##  ID     (Intercept) 0.97265  0.9862
##         Contrast1   0.16141  0.4018   -0.37
##  Item   (Intercept) 0.04943  0.2223
##         Match1      0.05900  0.2429   -0.08
## Number of obs: 5664, groups:  ID, 118; Item, 12
##
## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  1.77597    0.12151  14.615  < 2e-16 ***
## OI1         -0.05713    0.17274  -0.331    0.741
## OI2          0.25833    0.18245   1.416    0.157
## Contrast1   -0.46113    0.09272  -4.974 6.57e-07 ***
## Match1      -0.79667    0.08085  -9.854  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##           (Intr) OI1    OI2    Cntrs1
## OI1        0.006
## OI2        0.015  0.041
## Contrast1 -0.187 -0.148 -0.037
## Match1    -0.112 -0.013 -0.014  0.038
# converges and no singular fit
```

```
gm_en_reduced <- glmer(Correct ~ 1 +
                         (1 + Contrast | ID) +
                         (1 + Match | Item),
                       data = ES_data, family = "binomial",
                       control=glmerControl(optimizer="bobyqa",
optCtrl=list(maxfun=100000)))

anova(gm_en_reduced,gm_en03)
## Data: ES_data
## Models:
## gm_en_reduced: Correct ~ 1 + (1 + Contrast | ID) + (1 + Match | Item)
## gm_en03: Correct ~ OI + Contrast + Match + (1 + Contrast | ID) + (1 + Match |
Item)
##               npar    AIC    BIC  logLik deviance Chisq Df Pr(>Chisq)
## gm_en_reduced    7 5045.0 5091.5 -2515.5   5031.0
## gm_en03         11 5011.2 5084.3 -2494.6   4989.2 41.74  4  1.889e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
#significant improvement on reduced model
## adding in interactions
gm_en04 <- glmer(Correct ~ OI * Contrast + Match +
                   (1 + Contrast | ID) +
                   (1 + Match | Item),
                 data = ES_data, family = "binomial",
                 control=glmerControl(optimizer="bobyqa",
optCtrl=list(maxfun=100000)))
#print(summary(gm_en04))

anova(gm_en03, gm_en04)
## Data: ES_data
## Models:
## gm_en03: Correct ~ OI + Contrast + Match + (1 + Contrast | ID) + (1 + Match |
Item)
## gm_en04: Correct ~ OI * Contrast + Match + (1 + Contrast | ID) + (1 + Match |
Item)
##         npar    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
## gm_en03   11 5011.2 5084.3 -2494.6   4989.2
## gm_en04   13 5006.1 5092.4 -2490.1   4980.1 9.1434  2    0.01034 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
#significant improvement

#gm_en05 <- glmer(Correct ~ OI * Contrast * Match +
 #                 (1 + Contrast | ID) +
  #                (1 + Match | Item),
   #             data = ES_data, family = "binomial",
    #            control=glmerControl(optimizer="bobyqa",
optCtrl=list(maxfun=100000)))
#print(summary(gm_en05))
```

```
#anova(gm_en04, gm_en05)
#not a significant improvement

##add additional fixed effects

#gm_en06 <- glmer(Correct ~ OI * Contrast + Match + prop_distraction +
 #                  (1 + Contrast | ID) +
  #                 (1 + Match | Item),
   #              data = ES_data, family = "binomial",
    #             control=glmerControl(optimizer="bobyqa",
optCtrl=list(maxfun=100000)))
#print(summary(gm_en06))

#anova(gm_en04, gm_en06)
#marginal improvement

gm_en07 <- glmer(Correct ~ OI * Contrast + Match + age +
                   (1 + Contrast | ID) +
                   (1 + Match | Item),
                 data = ES_data, family = "binomial",
                 control=glmerControl(optimizer="bobyqa",
optCtrl=list(maxfun=100000)))

print(summary(gm_en07))
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula: Correct ~ OI * Contrast + Match + age + (1 + Contrast | ID) +
##     (1 + Match | Item)
##    Data: ES_data
## Control: glmerControl(optimizer = "bobyqa", optCtrl = list(maxfun = 1e+05))
##
##      AIC      BIC   logLik deviance df.resid
##   4998.8   5091.8  -2485.4   4970.8     5650
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -6.4263  0.1241  0.2989  0.4848  1.7819
##
## Random effects:
##  Groups Name        Variance Std.Dev. Corr
##  ID     (Intercept) 0.87054  0.9330
##         Contrast1   0.15919  0.3990   -0.35
##  Item   (Intercept) 0.01716  0.1310
##         Match1      0.06058  0.2461   -0.35
## Number of obs: 5664, groups:  ID, 118; Item, 12
##
## Fixed effects:
##                 Estimate Std. Error z value Pr(>|z|)
## (Intercept)     2.580606   0.282672   9.129  < 2e-16 ***
```

```
## OI1           -0.026358   0.118875  -0.222 0.824523
## OI2            0.269011   0.126540   2.126 0.033512 *
## Contrast1     -0.465613   0.071932  -6.473 9.61e-11 ***
## Match1        -0.798557   0.081698  -9.774  < 2e-16 ***
## age           -0.022898   0.007374  -3.105 0.001902 **
## OI1:Contrast1 -0.395582   0.109882  -3.600 0.000318 ***
## OI2:Contrast1 -0.111692   0.131843  -0.847 0.396908
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) OI1    OI2    Cntrs1 Match1 age    OI1:C1
## OI1          0.000
## OI2          0.012  0.036
## Contrast1   -0.071 -0.158 -0.037
## Match1      -0.085 -0.009 -0.016  0.038
## age         -0.927  0.004 -0.003 -0.025  0.013
## OI1:Cntrst1 -0.017 -0.098 -0.027  0.040  0.021  0.004
## OI2:Cntrst1 -0.004  0.092 -0.107 -0.057 -0.003 -0.001 -0.006

anova(gm_en04, gm_en07)
## Data: ES_data
## Models:
## gm_en04: Correct ~ OI * Contrast + Match + (1 + Contrast | ID) + (1 + Match |
Item)
## gm_en07: Correct ~ OI * Contrast + Match + age + (1 + Contrast | ID) + (1 +
Match | Item)
##         npar    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
## gm_en04   13 5006.1 5092.4 -2490.1   4980.1
## gm_en07   14 4998.8 5091.8 -2485.4   4970.8 9.3279  1   0.002257 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
#significant improvement

#gm_en08 <- glmer(Correct ~ OI * Contrast + Match + age + headphones +
 #                  (1 + Contrast | ID) +
  #                 (1 + Match | Item),
   #              data = ES_data, family = "binomial",
    #             control=glmerControl(optimizer="bobyqa",
optCtrl=list(maxfun=100000)))
#print(summary(gm_en08))

#anova(gm_en07, gm_en08)
#no significant improvement

#gm_en09 <- glmer(Correct ~ OI * Contrast + Match + age +
Participant.Device.Type +
 #                  (1 + Contrast | ID) +
  #                 (1 + Match | Item),
   #              data = ES_data, family = "binomial",
```

```
    #             control=glmerControl(optimizer="bobyqa",
optCtrl=list(maxfun=100000)))
#print(summary(gm_en09))

#anova(gm_en07, gm_en09)
#no significant improvement

#gm_en10 <- glmer(Correct ~ OI * Contrast + Match + education +
 #                 (1 + Contrast | ID) +
  #                (1 + Match | Item),
   #             data = ES_data, family = "binomial",
    #            control=glmerControl(optimizer="bobyqa",
optCtrl=list(maxfun=100000)))
#print(summary(gm_en10))

#anova(gm_en07, gm_en10)
#no significant improvement
```

To assess the contribution of each fixed effect to the fit of the model, further comparisons are performed using likelihood tests between the full final model and depleted versions.

```
#final model
#gm_en07 <- glmer(Correct ~ OI * Contrast + Match + age +
 #                 (1 + Contrast | ID) +
  #                (1 + Match | Item),
   #             data = ES_data, family = "binomial",
    #            control=glmerControl(optimizer="bobyqa",
optCtrl=list(maxfun=100000)))
#print(summary(gm_en07))

gm_enmatch_noOI <- glmer(Correct ~ Contrast + Match + age +
                   (1 + Contrast | ID) +
                   (1 + Match | Item),
                 data = ES_data, family = "binomial",
                 control=glmerControl(optimizer="bobyqa",
optCtrl=list(maxfun=100000)))

anova(gm_enmatch_noOI,gm_en07)
## Data: ES_data
## Models:
## gm_enmatch_noOI: Correct ~ Contrast + Match + age + (1 + Contrast | ID) + (1
+ Match | Item)
## gm_en07: Correct ~ OI * Contrast + Match + age + (1 + Contrast | ID) + (1 +
Match | Item)
##                 npar    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
## gm_enmatch_noOI   10 5001.9 5068.3 -2491.0   4981.9
## gm_en07           14 4998.8 5091.8 -2485.4   4970.8 11.151  4    0.02492 *
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
#significant

gm_enmatch_noContrast <- glmer(Correct ~ OI + Match + age +
                   (1 + Contrast | ID) +
                   (1 + Match | Item),
                 data = ES_data, family = "binomial",
                 control=glmerControl(optimizer="bobyqa",
optCtrl=list(maxfun=100000)))

anova(gm_enmatch_noContrast,gm_en07)
## Data: ES_data
## Models:
## gm_enmatch_noContrast: Correct ~ OI + Match + age + (1 + Contrast | ID) + (1
+ Match | Item)
## gm_en07: Correct ~ OI * Contrast + Match + age + (1 + Contrast | ID) + (1 +
Match | Item)
##                       npar    AIC    BIC  logLik deviance Chisq Df Pr(>Chisq)
## gm_enmatch_noContrast   11 5016.3 5089.3 -2497.1   4994.3
## gm_en07                 14 4998.8 5091.8 -2485.4   4970.8 23.51  3  3.161e-05
##
## gm_enmatch_noContrast
## gm_en07               ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
# highly significant

gm_enmatch_noMatch <- glmer(Correct ~ OI * Contrast + age +
                   (1 + Contrast | ID) +
                   (1 + Match | Item),
                 data = ES_data, family = "binomial",
                 control=glmerControl(optimizer="bobyqa",
optCtrl=list(maxfun=100000)))
anova(gm_enmatch_noMatch,gm_en07)
## Data: ES_data
## Models:
## gm_enmatch_noMatch: Correct ~ OI * Contrast + age + (1 + Contrast | ID) + (1
+ Match | Item)
## gm_en07: Correct ~ OI * Contrast + Match + age + (1 + Contrast | ID) + (1 +
Match | Item)
##                    npar    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
## gm_enmatch_noMatch   13 5022.9 5109.3 -2498.5   4996.9
## gm_en07              14 4998.8 5091.8 -2485.4   4970.8 26.145  1  3.167e-07
***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
# highly significant

gm_enmatch_noage <- glmer(Correct ~ OI * Contrast + Match +
                   (1 + Contrast | ID) +
```

```
                   (1 + Match | Item),
                 data = ES_data, family = "binomial",
                 control=glmerControl(optimizer="bobyqa",
optCtrl=list(maxfun=100000)))

anova(gm_enmatch_noage,gm_en07)
## Data: ES_data
## Models:
## gm_enmatch_noage: Correct ~ OI * Contrast + Match + (1 + Contrast | ID) + (1
+ Match | Item)
## gm_en07: Correct ~ OI * Contrast + Match + age + (1 + Contrast | ID) + (1 +
Match | Item)
##                  npar    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
## gm_enmatch_noage   13 5006.1 5092.4 -2490.1   4980.1
## gm_en07            14 4998.8 5091.8 -2485.4   4970.8 9.3279  1   0.002257 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
#significant

gm_enmatch_noInteract <- glmer(Correct ~ OI + Contrast + Match + age +
                   (1 + Contrast | ID) +
                   (1 + Match | Item),
                 data = ES_data, family = "binomial",
                 control=glmerControl(optimizer="bobyqa",
optCtrl=list(maxfun=100000)))

anova(gm_enmatch_noInteract,gm_en07)
## Data: ES_data
## Models:
## gm_enmatch_noInteract: Correct ~ OI + Contrast + Match + age + (1 + Contrast
| ID) + (1 + Match | Item)
## gm_en07: Correct ~ OI * Contrast + Match + age + (1 + Contrast | ID) + (1 +
Match | Item)
##                       npar    AIC    BIC  logLik deviance  Chisq Df
Pr(>Chisq)
## gm_enmatch_noInteract   12 5003.9 5083.6 -2490.0   4979.9
## gm_en07                 14 4998.8 5091.8 -2485.4   4970.8 9.1538  2
0.01029
##
## gm_enmatch_noInteract
## gm_en07               *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
# significant
```

Appendix XI: Coder guidelines for Q6

The different categories are outlined in the tables below, accompanied by definitions and examples. There are two coding guides, one for the main categories of overall orthographic input (OI) influence and another for the categories specifying the influence of OI. There are 5 positive categories, containing 5 further sub-categories, as well as 3 negative/ no influence categories. Examples for how codes can be applied are then provided in another table. At the end of the document, there is a random sample of 20% of participant responses which should be coded according to the coding guidelines provided. The recommended steps are:

1. Code all complete phrases for the overall evaluation of OI influence i.e. positive or negative/ no influence.
2. Discard responses that are too vague to be coded or are not directly related to the category definition or research question.
3. Go back through material and follow the second coding guide to apply the more specific categories within each main type.
4. Make notes of any categories that you feel do not fit well or responses that may require the formulation of a new or different category.

Of course, if you have any questions, don't hesitate to ask!

Coding Arabic responses

Table XI.1: L1 Arabic coding guide for overall evaluation of OI influence during the task

Type	Definition	Examples
Positive influence (OI helps)	Seeing the written form was helpful or easier. It enabled, aided or meant you could/were able to do something that facilitated learning, understanding or memory. This could be general or specific to English or Arabic script input **Arabic keywords** أصبح/ جعلها أسهل/ أسرع, يساعدني	"Easier to remember" "It's very helpful when i see the word written." "اذا كانت الكلمة مكتوبة بالانجليزية اسهل"
Negative/Neutral influence (OI does not help)	Seeing the written form was unhelpful, introduced added difficulty or influence that inhibited learning, understanding or memory. Alternatively, seeing the written form didn't make a noticeable difference. Even if there was no explicit negative outcome, it did not help. **Arabic keywords** لا يوجد فرق, كانت اصعب/ يشتت	"Arabic written words were more of a distraction." "I read them when I noticed them but it didn't seem to help" "الكلمات المكتوبة كانت اصعب"

Table XI.2: L1 Arabic coding guide for specific perceptions of OI influence

Type	Category	Definition	Examples
Positive (OI helps…)	Learning (broadly)	Facilitates learning in general. Makes learning easier/ better/ faster. Does not specify further. **Arabic keywords** تعلمها/ استيعابها اسهل/ اسرع / أكثر	"Yes, it makes learning easier." "التعلم اسرع بوجود الكلمات المكتوبة"
	Remembering (broadly)	Facilitates memory/ recall/ retention of words or objects. Does not specify beyond general memory/ memorisation aid. **Arabic keywords** يسهل التذكر / الحفظ	"Much easier to remember" "رؤية الكلمات العربية سهل تذكرها بالنسبة لي اكثر من الانجليزية"
	Focusing on first letter	Specific to initial letter and not sound; use of first letter as a memory aid. **Arabic keywords** تذكر الحرف الأول	"I think it was easier to remember the first letter"
	Connect audio and visual input (broadly)	Useful to connect/ associate/ match / link the different types or sources of input; references ideas that align with multimodal input and dual-coding; does not specify whether association is between spelling, sound or image. May reference learning and memory. **Arabic keywords** ربط البصر والصوت (معًا/ سويا)	"Seeing English spelling helped to make connections between what I was hearing and what I was seeing." "الاستماع والقراءة معا تجعل من التعلم أسهل"
	Visualise words	Specifically mentions visualising, imagining or picturing the word in the mind; implies a focus on the whole word and doesn't mention the image or object meaning. **Arabic keywords**	"I could visualise what I was hearing"

		تخيل شكل الكلمة	"نعم فقد ساعدني في تذكر شكل الكلمة كثيرا سواء الكلمة بالعربية او الانجليزية فأنا أتخيلها"
	Map sounds to letters	Specifically mentions the connection between spelling and sounds; focus on individual letters and sounds rather than the whole word; may reference active processes of spelling out or reading along with sound. **Arabic keywords** ربط/ تطابق الحروف والصوت	"In English it helped because it gave me something else to remember and map sounds onto." "رؤية الكلمة الإنجليزية يساعدني أحيانا في استيعاب المسموع وخاصة أن الكلمات الإنجليزية تطابق حروفها الصوت المسموع"
	Associate word and image	Specifically mentions the connection between the (spelling of the) word and the image/ object/ item. **Arabic keywords** ربط/ تطابق الصورة والصوت/ الكلمة	"The English words helped my brain link the images to the sound better." "كانت الكلمات المكتوبة باللغة الإنجليزية أدعى لتذكرها والربط بينها وبين الصورة"
	Clarify what I heard	know/ understand/ comprehend spoken form/ pronunciation; check hearing was correct; identify sounds that were missed with audio only. **Arabic keywords** تأكيد/ معرفة للكلمة المسموعة/ النطق الصحيح, لفهم أفضل	"Yes, in English because it helped me to clarify if I had heard the spoken pronunciation correctly." "الكلمات المكتوبة مصدر تأكيد للكلمة المسموعة"
	Differentiate similar sounds/words	Distinguish/ clarify/ differentiate/ separate/ learn words and sounds that are similar. May mention the first letter but in the	"I think it aided the distinction between similar sounding words"

		context of differentiating similar sounds, rather than as a memory aid. **Arabic keywords** التفريق/ التمييز بين الكلمات/ الأصوات المتشابهة	"الفرق بدى واضحًا في الكلمات المتشابهه مثل Fan Van "
	Rehearse word	Practice the word out loud or in their mind. **Arabic keywords** تدرب/ كرر على الكلمات/ الأصوات	"Seeing them helped me rehearse the word in my head."
Negative/ Neutral (OI didn't help because it was…)	Distracting (from sounds)	Written words were a distraction in general; related to cognitive overload or inconsistent mappings between sound and spelling. **Arabic keywords** يشتت	"The Arabic words put with the picture made it so that I felt there was more to take in, so distracted from learning the sound" "في الغالب لا يحدث فرق بل ربما يشتت قليلا النظر الى الكلمة المكتوبة والصورة والاستماع"
	Unfamiliar	Written forms were written in an unfamiliar script, so unable to read or distinguish the words. **Arabic keywords** غير مألوفة, لا أستطيع قراءتها	"The Arabic spelling didn't make any difference because I do not know how to read it."
	Better to focus on sounds	Ignored written forms; chose to pay attention/ concentrate/ focus on the sound and/or image **Arabic keywords** أركز على الصوت, كان سماع الكلمات أفضل/أسهل	"I focus more on hearing them out loud and trying to match the sound to the colours and shapes." "لا، أركز فقط على الصوت"

Table XI.3: L1 Arabic example of coding approach

ID	When learning the new words, did it make a difference seeing the written words? In what way?	codes
5359091	ربط الصور مع الكلمات يجعل عمليه التذكر اسهل	Positive - associate word and image
5462535	to the remember of the words	Positive - remembering
5097676	يجعل التذكر اسرع	Positive - remembering
5015277	yes ,i prefer the word written while listening to it , that's how i memorize the word twice written and spoken	Positive - connect audio and visual input
5357676	نعم فقد ساعدني في تذكر شكل الكلمة كثيرا سواء الكلمة بالعربية او الانجليزية فأنا أتخيلها واتذكر لو حتى حرفا واحدا	Positive - visualise words

Table XI.4: L1 Arabic 20% random sample for 2nd coder

ID	When learning the new words, did it make a difference seeing the written words? In what way?	codes
5189512	It made it easier to remember the words	
5089630	yes, it makes it easier to remembr the word because you saw its spelling	
4986648	differentiate betweet the V and F	
5201938	to remember the word	
4986647	نعم. انا انسانة بصرية	
5370466	لا، أركز فقط على الصوت	
5034149	it's very helpful when I see the word written	
5027661	yes made a difference , but not that much	
5029568	نعم إلى مدى مهم في تنشيط الذاكرة والحفظ	
5011969	نعم أحدث فرقا إلي مدى كبير	
5359752	نعم،كتابة الكلمة باللغة الانجليزية تجعل تعلمها اسهل	

5339892	نعم، أصبحت الفروق واضحة بين F و V. كما كانت القراءة والاستماع في آن وحدا وسيلة للتأكيد على حفظ الكلمة.	
5499774	Yes, How much it helps to learn.	
5338948	جعل تذكرها اسهل	
4986649	written words are much easier to remember and to associate with the right item	
5450539	في الغالب لا يحدث فرق بل ربما يشتت قليلا النظر الى الكلمة المكتوبة والصورة والاستماع	
5229722	As some of the words sound similar, seeing it wtitten helped me distinguish between them when hearing them so I would say it helped a lot in that case, it also helped with other cases as linking the image and the sound with a written word made it easier to remember.	
5452186	helped me remember the words as I am a visual person	
5027570	i can remember the word spelling, so each time i see the picture i remember the word, but this didn't happen to me in all the pictures only in some of them.	
5037046	الكلمات المكتوبة مصدر تأكيد للكلمة المسموعة	
5038591	التعلم اسرع بوجود الكلمات المكتوبة	
5031281	watching the written words make learning easier as I can link the image with the written word	
5082295	نعم ، لمدى كثير جدا بالمكتوبه استطيع معرفه النطق الصحيح	
5335791	رؤية الكلمات العربية سهل تذكرها بالنسبة لي اكثر من الانجليزية	
5442543	بعض الكلمات لم ألاحظ أنها تبدأ بصوتين متشابهين وبما أن هذه الأصوات غائبة في لغتي في الغالب لن أنتبه لها عند v and f مثلا سماعها فقط فالإملاء يساعد في هذه الحالة	

Coding English responses

Every unit of analysis was coded for general type of influence (positive, negative/no influence) first, and then further categorised according to any additional details about the specific ways in which seeing written forms influenced their ability to process, remember or learn the target items in the study. The positive responses all implied an enhancement of learning with written forms, thus categories all follow on from the phrase “OI helps…” to capture the facilitatory effects. The negative and neutral responses tended to include more information about what it was about the written forms that made them unhelpful or unnecessary for their learning, thus categories follow on from the phrase “OI did not help because it was…”. These categories can apply broadly or in specific reference to English or Arabic script.

Table XI.5: L1 English coding guide for overall evaluation of OI influence during the task

Type	**Definition**	**Examples**
Positive influence (OI helps)	Seeing the written form was helpful or easier. It enabled, aided or meant you could/were able to do something that facilitated learning, understanding or memory. This could be general or specific to English or Arabic script input	“Easier to remember” “I could visualise what I was hearing later.”
Negative/Neutral influence (OI does not help)	Seeing the written form was unhelpful, introduced added difficulty or influence that inhibited learning, understanding or memory. Alternatively, seeing the written form didn’t make a noticeable difference. Even if there was no explicit negative outcome, it did not help.	“Arabic written words were more of a distraction.” “I read them when I noticed them but it didn’t seem to help”

Table XI.6: L1 English coding guide for specific perceptions of OI influence

Type	Category	Definition	Examples
Positive (OI helps…)	Learning (broadly)	Facilitates learning in general. Makes learning easier/ better/ faster. Does not specify further.	"Yes, it makes learning easier."
	Remembering (broadly)	Facilitates memory/ recall/ retention of words or objects. Does not specify beyond general memory aid.	"Much easier to remember"
	Focusing on first letter	Specific to initial letter and not sound; use of first letter as a memory aid.	"I think it was easier to remember the first letter"
	Connect audio and visual input (broadly)	Useful to connect/ associate/ match / link the different types or sources of input; references ideas that align with multimodal input and dual-coding; does not specify whether association is between spelling, sound or image.	"Seeing English spelling helped to make connections between what I was hearing and what I was seeing."
	Visualise words	Specifically mentions visualising or picturing the word; implies a focus on the whole word and doesn't mention the image or object meaning.	"I could visualise what I was hearing"
	Map sounds to letters	Specifically mentions the connection between spelling and sounds; focus on sounds rather than the whole word; may reference active processes of spelling out or reading along with sound.	"In English it helped because it gave me something else to remember and map sounds onto."
	Associate word and image	Specifically mentions the connection between the (spelling of the) word and the image/ object.	"The English words helped my brain link the images to the sound better."

	Clarify what I heard Differentiate similar sounds/words	know/ understand/ comprehend spoken form/ pronunciation; check hearing was correct; identify sounds that were missed with audio only. Distinguish/ clarify/ differentiate/ separate/ learn words and sounds that are similar. May mention the first letter but in the context of differentiating similar sounds, rather than as a memory aid.	"Yes, in English because it helped me to clarify if I had heard the spoken pronunciation correctly." "I think it aided the distinction between similar sounding words"
	Rehearse word	Practise the word out loud or in their mind.	"I would try to say them if they were written in English."
Negative/ Neutral (OI didn't help because it was…)	Distracting (from sounds)	Written words were a distraction in general; related to cognitive overload or inconsistent mappings between sound and spelling.	"The Arabic words put with the picture made it so that I felt there was more to take in, so distracted from learning the sound"
	Unfamiliar	Written forms were written in an unfamiliar script, so unable to read or distinguish the words.	"The Arabic spelling didn't make any difference because I do not know how to read it."
	Better to focus on sounds	Ignored written forms; chose to pay attention/ concentrate/ focus on the sound and/or image	"I focus more on hearing them out loud and trying to match the sound to the colours and shapes."

Table XI.7: L1 English example of coding approach

ID	When learning the new words, did it make a difference seeing the written words? In what way?	codes
5305661	Easier to visualise	Positive - Visualise words
5308502	i think it aided the distinction between similar sounding words	Positive - Differentiate similar sounds/ words
5308657	Yes, in English because it helped me to clarify if I had heard the spoken pronunciation correctly. I was often surprised at the spelling (e.g. "nadus" spelt nad-I-s).	Positive - Clarify what I heard
5308960	In english, enabled me to spell it out in my head	Positive - Map sounds to letters
5312548	I found English written words slightly helpful, while Arabic written words were more of a distraction.	Positive Negative - Distracting

Table XI.8: L1 English 20% random sample for 2nd coder

ID	When learning the new words, did it make a difference seeing the written words? In what way?	codes
5692078	Helpful when in English	
5676564	Yes I was able to see be sure of whether the speaker said m/n or v/f as they can sound quite similar. i was also able to remember the spelling and repeat. hearing the word aloud from myself helped me	
5341833	Helped with initial consonent (e.g. v/f, m/n)	
5356537	much easier in english, and seeing the spelling helped in remembering	
5765135	it helped put some letters to a sound, this made it easier because it was remembering more distinct qualities related to the sound that was played	
5374964	yes, i think it helped me to remember	
5418540	Seeing the word made it easier to understand what the word was, because even if you misheard the word, the spelling was there so reinforce what the word was. It also meant you heard the word and saw the word so you had two chances to learn the word.	

5664502	For English, being able to distinguish minimal pairs, e.g., m/n and f/v	
5755297	Yes, I am a visual learner, so being able to connect the English spelling of the words with the images helped me differentiate the images with similar sounds	
5312160	It makes a difference to me in general because I can see the sounds I am unable to pick up by just listening.	
5304273	yes - easier to help remember when in english	
5909763	Yes - helpful to gain a sense of pronunciation and structure	
5755301	the writing helped to clarify which initial sound the word had (v or f, m or n, etc)	
5499420	Yes, seeing them helped me rehearse the word in my head. It also confirmed the sounds that the speaker was making (e.g. v versus f)	
5343885	When they were written in English it made it easier to remember for me	
5909329	reinforces the sound of the word	
5762777	it was easier to differentiate the consonant sounds after seeing the spelling in english	
5328002	Yes, easier to associate the word with the object when you can see the spelling	
5759735	yes, it helped with pronouncing	
5691858	Yes. It was easier to remember the object described.	
5671003	I wasn't really paying attention to the written words at first so I thought it made no difference, but on some level I must have been paying attention because when they were no longer included, it became a little more difficult to hear what letters were being said.	
5706491	In English it helped because I could begin to match them together when I saw them repeatedly over time.	
5417849	Yes, helped to know which letter each word started with.	
5649032	It helped to see the English versions written out to match sounds of pronunciation with what I was hearing. But without words I focused more on the sounds and being able to differentiate eg similar prefixes.	

Appendix XII: Coder guidelines for Q8

The different categories are outlined in the tables below, accompanied by definitions and examples. The main affirmative or negative categories of OI importance are predetermined by participant answers YES-NO response to the initial question above. The open responses, corresponding to the 'why?' part of the question is then to be coded according to the coding guide provided below. There are 7 positive categories, one of which contains 3 sub-categories, then 4 negative categories. Examples for how codes can be applied are then provided in another table. At the end of the document, there is a random sample of 20% of participant responses which should be coded according to the coding guidelines provided.

The recommended steps are:

1. Code all complete phrases for the overall evaluation of OI importance i.e. early exposure to OI is (not) important.
2. Go back through material and follow the coding guide to apply the more specific categories within each main type.
3. Discard responses that are too vague to be coded or are not directly related to the category definition or research question.
4. Make notes of any categories that you feel do not fit well or responses that may require the formulation of a new or different category.

Of course, if you have any questions, don't hesitate to ask!

Coding Arabic responses

Every unit of analysis was coded for general belief (early exposure to OI is (not) important), and then further categorised according to additional rationale for this belief, including specific ways that seeing written forms impacts the memory, processing and general learning of words in a new language. The affirmative responses all follow on from the phrase "Early OI is important for…", whereas the negative responses follow on from the phrase "Early OI is not important because…". L1 Arabic-speaking participants could choose whether to respond in English or Arabic and the original, untranslated and unchanged responses have been provided for coding. All categories are English, but keywords and examples have been provided in Arabic.

Table XII.1: L1 Arabic coding guide for specific beliefs around OI importance

Type	Category	Definition	Examples
Early OI is important for…	Remembering (broadly)	The additional visual cue from written input helps remember/recall the words, in general. Does not specify further. **Arabic keywords** تذكر	"I think it helps me remember them" "Additional aid to memory" "يسهل تذكرها"
	Visualisation	Finds visualising, imagining or picturing the (shape of the) word is helpful and may refer to memory and learning. Does not specify further. **Arabic keywords** تخيل/ تذكر شكل الكلمة, في مخيلتي/ في ذهني	"Helps me to see the word and visualise it." "حتى اتخيل شكل الكلمة"
	Visual association	The written input facilitates connections/ associations/ links between verbal and visual input, as well as linking to existing linguistic knowledge. Could function as a cue/ / hint/ clue/ reminder. Not specific to individual sounds and letters, but rather broader connections between word, spelling and meaning. May contain ideas related to dual-coding, as well as memory and learning. **Arabic keywords** ربط الكلمة المكتوبة مع الصوت أو المعنى / النطق و الكتابة معا	"You can make stronger connections between the word, its spelling and the sound and the concept it represents." "لكي اربط في ذهني الكلمة المكتوبة مع الصوت ... احفظ النطق و الكتابة معا في نفس الوقت
	Memorisation	Connect the written input to the process of memorisation/ rote learning/ learning by heart. May make it an easier/ better process. Does not specify further **Arabic keywords** الحفظ	"In order to memorise that word." "لكي يسهل حفظها"

	Clarifying and consolidating new language	Written input helps to "know" and feel sure about the correct form of the word. It helps recognize and embed new language in the mind **Arabic keywords** تثبيت / ترسيخه/ ترتكز في المخ, معرفتها	"In order to create a clear mental representation in my mind. "هو بمثابة تثبيت للكلمة"
	Correct comprehension	Specifically helps hear/ listen/ understand/ comprehend/ realise/ confirm the correct form of the word. Avoid mishearing spoken form/ pronunciation, especially difficult to perceive sounds. Focus on perception and receptive language processing, not production. **Arabic keywords** لتميي/ تفريق بين كلمة أخرى شبيهة, لافهم, للتأكد ما استمعت إليه	"It prevents you from learning it incorrectly incase you have misheard the pronunciation" "للتأكد من صحة ما استمعت إليه"
	Correct production	Specifically helps say/ pronounce/ write/ make/ a word or sound correctly. Avoids frustrations or insecurity around using the language accurately. Focus on usage and productive language processing, not perception. **Arabic keywords** النطق الصحيح,استخدامها في التحدث والكتابة	"You can also make sure you are pronouncing it correctly if you see it." "لأعرف نطق الكلمة." "لمعرفة كيفية كتابتها"
	Decoding sounds and letters	Matching/ looking for patterns/ breaking down/ sounding out sounds and letters. May mention learning phonetically, which is interpreted to mean learning through one-to-one sound-letter mappings, as well as supporting literacy acquisition or reading. May mention vowels and diacritics in Arabic script. **Arabic keywords** املاء وصوتها, الأنماط بين الحروف والأصوات, حركات التشكيل	"It helps for a learner to recognise patterns within the spelling system and how they interact with the sounds. "حتى اربط بين املاء الكلمة وصوتها"

	Learning style	Identify their personal learning style, particularly being visual or textual learner. May state this as a preference rather than learner identity. **Arabic keywords** ذاكرتي صورية, متعلم بصري	“I am a more visual learner than an auditory learner.” “لأني بصرية”
	Speed of learning	Written input speeds up the process of learning, remembering, memorising, absorbing the new words. **Arabic keywords** حفظ/ تعلم بشكل أسرع, بسرعة	“I can recognise it and remember it faster.” "تثبت الكلمة اسرع "
Early OI is not important (because…)	Sound is more important/beneficial	Places importance on knowing the sound(s) well. This is the best point of focus. It can help know and remember the phonological form of the word as well as rehearse it orally. **Arabic keywords** التركيز على الصوت, الصوت أكثر أهمية	“If the focus is purely on sound, hearing and then repeating orally is more beneficial in consistently replicating and identifying the sound/word.”
	It is too much to process	The additional visual input is too much to process with the sounds as well. It interferes/ distracts from sounds or generally overloading. It introduces unnecessary elements that get in the way of learning and memory (cognitive load). **Arabic keywords** اتشتت عن استماع, أكثر من اللازم	“I think it interferes with my listening to see the word.” “لانه يصعب علي المهمة في حفظ الاملاء للاحرف وحفظ الكلمة المستمعة”
	Spelling doesn’t reflect pronunciation	Pronunciation does not consistently match the spelling of words. Spelling can be confusing/ misleading/ unreliable. May refer to spelling-sound correspondences in specific languages (orthographic depth). **Arabic keywords**	“Spelling in English can be confusing.” “Words can be pronounced differently to how they look.”

		الأنماط لا تتطابق مع النطق	"اللغة الإنجليزية أصوات غير متوفرة في حروف العربية"
	Of previous language learning	Reference previous experiences of language language as evidence for the importance of written input. May include general ideas and observations about language learning or personal preferences and strategies. **Arabic keywords** أفضل أن أتعلم من خلال التحدث والاستماع, الطريقة المألوفة للتعلم	"I prefer to learn languages by speaking and hearing them, not writing" "انا شخصية سمعية"

Table XII.2: L1 Arabic example of coding approach

ID	In general, when learning new words do you think it is important to see the spelling the first time you hear it? Why?		codes
5370516	yes	كما أشرت في الإجابة على سؤال سابق،، أولا: الكلمات العربية تحتاج إلى حركات التشكيل لتؤدي نفس الغرض الذي تؤديه الكلمات الإنجليزية للصوت نفسه، وثانية في اللغة الإنجليزية أصوات غير متوفرة في حروف العربية، (f and v) فمثلا الأحرف العربية لا تفرق بين صوتي	Decoding sounds and letters Spelling doesn't reflect pronunciation
5082294	yes	to remember it	remembering
5201553	yes	This way you remember it better.	remembering
5336722	yes	To realize how it's pronounced later when I see the word.	Correct production
5029572	yes	احفظها بشكل اسرع	Speed of learning
5365353	no	sound moreimportant	Sound is more important/beneficial

Table XII.3: L1 Arabic 20% random sample for 2nd coder

ID	In general, when learning new words do you think it is important to see the spelling the first time you hear it? Why?		codes
5229722	yes	This way I got more information which means more links to the word to easily recall it.	
4986648	yes	make sure how it is prounanced	
5359092	yes	افهم الكلمة اكثر	
5341797	yes	حتى أعرف النطق الصحيح، لأنني أحيانا لا أفرق بين f و v. فرؤية الكلمات مكتوبة يساعد على التعلم	
5499647	no	لكي لا اتشتت بالتركيز على املاء تلك الكلمة	
5027737	yes	Seeing the spelling helps me remember the word better.	
5010884	yes	لأن العين تحفظ ايضا ولعدم القدرة على تمييز بعض الاحرف عند سماعها فصعبت التفريق بين نطق بعض الكلمات	
5034149	yes	to help me connect the voice with the letters	

5337282	yes	its help me to draw a vision on my mind and link between the word and the shape	
4986647	yes	حتى اتخيل شكل الكلمة	
5029580	yes	reading the spelling of a word makes it easier to learn	
5338811	yes	حتى ترتكز في المخ	
5501560	yes	to memorize it in better way	
4986649	yes	notice the difference between v and f better	
5331395	yes	لكي افهمها جيداً	
5384016	yes	لتثبت في مخيلتي	
5031282	yes	يسهل تذكرها	
5378097	yes	لكي يسهم البصر في حفظ الكلمة	
5003023	yes	لكي اربط في ذهني الكلمة المكتوبة مع الصوت ... احفظ النطق و الكتابة معا في نفس الوقت	
5003012	yes	تساعدني على ربطها بالصورة و تخزينها في الذاكرة يكون اسهل و اسرع	
5037277	yes	in order to memorize that word	
5182923	yes	to know how to spell it	
5212731	no	ممكن أن أتشتت عن استماع الكلمة بالشكل الصحيح وأركز على شكها فقط	
5357676	yes	يساعدني جدا في تذكرها وربطها بكلمة في لغتي الام ولو برابط مضحك خاص بي انا .. لا اخبره لأحد	
5029581	yes	to pronunce it correctly	

Coding English responses

Every unit of analysis was coded for general belief (early exposure to OI is (not) important), and then further categorised according to additional rationale for this belief, including specific ways that seeing written forms impacts the memory, processing and general learning of words in a new language. The affirmative responses all follow on from the phrase “Early OI is important for…”, whereas the negative responses follow on from the phrase “Early OI is not important because…”.

Table XII.4: L1 English coding guide for specific beliefs around OI importance

Type	**Category**	**Definition**	**Examples**
Early OI is important for…	Remembering (broadly)	The additional visual cue from written input helps remember/recall the words, in general. Does not specify further.	"I think it helps me remember them" "Additional aid to memory"
	Visualisation	Finds visualising or picturing the word helpful and may refer to memory and learning. Does not specify further.	"Helps me to see the word and visualise it."
	Visual association	The written input facilitates connections and associations between verbal and visual input, as well as linking to existing linguistic knowledge. Not specific to individual sounds and letters, but rather broader connections between word, spelling and meaning. May refer to memory and learning.	"You can make stronger connections between the word, its spelling and the sound and the concept it represents." "Easier to remember the word and place it with the object in my memory."
	Memorisation	Connect the written input to the process of memorisation/ rote learning/ learning by heart. May make it an easier/ better process.	"In order to memorise that word."
	Clarifying and consolidating new language Correct comprehension	Written input helps to "know"/ be familiar with/ feel sure about the correct form of the word. It helps embed new language in the mind Specifically helps hear/ listen/ understand/ comprehend/ realise/ confirm the correct form of the word. Avoid mishearing spoken form/ pronunciation, especially difficult to perceive sounds. Focus on perception and receptive language processing, not production.	"It helps me understand what the word means." "It prevents you from learning it incorrectly incase you have misheard the pronunciation" "Solidifies understanding of hearing the word, like a confirmation."

	Correct production Decoding sounds and letters	Specifically helps say/ pronounce/ write/ make/ a word or sound correctly. Avoids frustrations or insecurity around using the language accurately. Focus on usage and productive language processing, not perception. Matching/ looking for patterns/ breaking down/ sounding out sounds and letters. May mention learning phonetically, which is interpreted to mean learning through one-to-one sound-letter mappings, as well as supporting literacy acquisition or reading.	"You can also make sure you are pronouncing it correctly if you see it." "Learning the spelling of a word allows you to use it in writing." "It helps for a learner to recognise patterns within the spelling system and how they interact with the sounds. "It is helpful to see the sound matched up with the letters."
	Learning style	Identify their personal learning style, particularly being visual or textual learner. May state this as a preference rather than learner identity.	"I am a more visual learner than an auditory learner." "I guess I maybe learn visually better than hearing the words."
	Speed of learning	Written input speeds up the process of learning, remembering, memorising, absorbing the new words.	"I can recognise it and remember it faster."
Early OI is not important (because…)	Sound is more important/ beneficial	Places importance on knowing the sound(s) well. This is the best point of focus. It can help know and remember the phonological form of the word as well as rehearse it orally.	"If the focus is purely on sound, hearing and then repeating orally is more beneficial in consistently replicating and identifying the sound/word."

	It is too much to process	The additional visual input is too much to process with the sounds as well. It interferes/ distracts from sounds or generally overloading. It introduces unnecessary elements that get in the way of learning and memory (cognitive load).	"I think it interferes with my listening to see the word." "Hearing is most helpful first. Written and hearing the word is more to process."
	Spelling doesn't reflect pronunciation	Pronunciation does not consistently match the spelling of words. Spelling can be confusing/ misleading/ unreliable. May refer to spelling-sound correspondences in specific languages (orthographic depth).	"Spelling in English can be confusing." "Words can be pronounced differently to how they look."
	Of previous (language) learning	Reference previous experiences of (language) learning where written input did not prove essential or beneficial. May include general ideas and observations about language learning or personal preferences and strategies.	"I prefer to learn languages by speaking and hearing them, not writing"

Table XII.5: L1 English example of coding approach

ID	In general, when learning new words do you think it is important to see the spelling the first time you hear it? Why?		codes
5332693	yes	Visual learner, my memory seems quite dependent on having something in front of me to help with learning	Learning style Remembering
5762777	yes	better to know which consonant is being spoken right away to remember it correctly	Correct comprehension
5756085	yes	I am a text-based learner, I believe	Learning style
5621683	no	I can memorize a word and a meaning pretty quickly. The spelling adds another layer of memorization.	Too much to process
5719579	yes	That way you know for sure what the sounds are that make up the word, and that there isn't any pronunciation variation	Correct comprehension
5521508	no	Spelling rarely reflects pronunciation in English and in the languages I have learnt.	Spelling doesn't reflect pronunciation

Table XII.6: L1 English 20% random sample for 2nd coder

ID	In general, when learning new words do you think it is important to see the spelling the first time you hear it? Why?		codes
5499421	yes	I think I find it quite difficult to make the correct sounds when speaking, without first seeing the word written down (assuming it is written in characters I am familiar with)	
5440728	no	Sometimes the spelling and the pronunciation differ and I'd rather hear the word first and then read it.	
5356537	yes	to ensure dont mishear, esoecially with similar letters, like v and f or m and n	
5356536	no	speaking and listening is the first skills in a new language	
5663887	yes	It makes it easier for me to learn, I am a pretty visual learner	
5759398	yes	to correct hearing errors	
5384032	yes	i find it easier to read the word first before looking at the picture	

5332522	yes	it helps me sound it out loud, which helps me with pronunciation and word recognition, as well as with recall	
5751031	no	For me, I find oral fluency much easier to reach than written. I prefer to learn languages by speaking and hearing them, not writing (which I have done as well, with dead languages).	
5587514	no	so many english words are spelt in different ways that dont always make it easy on first appearance	
5732647	yes	This clarifies any misconception with sounds, especially when the words are similar	
5759735	no	hearing is most helpful first. written and hering the word is more to process	
5336201	yes	because you can then visualise the word as well, and maybe notice patterns in spellings if it is similar to a word you already know. You can also make sure you are pronouncing it correctly if you see it. (if you can read)	
5335169	yes	Because some letters sound similar, it re-enforces you are hearing the correct sound	
5308657	yes	Because I'm quite a visual and textual learner and slightly deaf, so I rely more on what I see than what I hear.	
5586643	yes	soliifies understanding of hearing the word, like a confirmation	
5914367	no	Easier to concentrate on one aspect of the word first. ie just hearing it or just seeing it	
6029761	yes	If you are a very visual learner, you can picture the word spelling when you say it	
5627982	yes	Helps embed in uour brain	
5338550	yes	to confirm youve heard the sounds correctly	
5676564	yes	to match the letters to the sounds	
5312548	yes	Learning the spelling of a word allows you to use it in writing. I find it frustrating being unsure about the spelling of a word when I want to use it in writing.	
5592006	yes	spellings provide a second sense to link the memory of the word to (sight as well as sound) and can give you clues as to keaning which can also help memory.	
5503474	yes	It can give an indication of which sound you are lisening for in English - though not always	

Appendix XIII: Intercoder reliability calculations Q6 and Q8

This document outlines the procedure to calculate Krippendorff's alpha, to assess intercoder agreement between three coders, for both L1 English and L1 Arabic data in R (irr package). Relevant code and output are included below, adapted from Rmd files. First, it is necessary to ensure that the data is in the correct long form before loading into R (see table XIII.1). This is then reshaped to create new data frames for each variable, with the content by coder (i.e. variable 1 = general_code, variable 2 = specific_code). This is performed with the data for both Q6 and Q8, and both L1 groups.

Table XIII.1 Example of data format before running R script

Participant_ID	Response	Content_ID	Coder_ID	General_code	Specific_code
1	xxxxx	1	A	Positive	Memory
2	xxxxx	2	A	Negative	Distraction
1	xxxxx	1	B	Positive	Visualisation
2	xxxxx	2	B	Negative	Distraction
1	xxxxx	1	C	Positive	Memory
2	xxxxx	2	C	Negative	Distraction

```
#English coders general perception of OI input
en_general_2.1 <- select(irr_2.1en, content_ID, coder_ID, general_code)
en_general_2.1 <- pivot_wider(en_general_2.1, id_cols = coder_ID,
names_from = content_ID, values_from = general_code)
en_general_2.1 <- select(en_general_2.1, -coder_ID)

#English coders specific perception of OI input
en_specific_2.1 <- select(irr_2.1en, content_ID, coder_ID,
specific_code)
en_specific_2.1 <- pivot_wider(en_specific_2.1, id_cols = coder_ID,
names_from = content_ID, values_from = specific_code)
en_specific_2.1 <- select(en_specific_2.1, -coder_ID)

#English coders specific belief of OI importance
en_specific_2.2 <- select(irr_2.2en, content_ID, coder_ID,
specific_code)
en_specific_2.2 <- pivot_wider(en_specific_2.2, id_cols = coder_ID,
names_from = content_ID, values_from = specific_code)
en_specific_2.2 <- select(en_specific_2.2, -coder_ID)
```

```
#Arabic coders general perception of OI input
ar_general_2.1 <- select(irr_2.1ar, content_ID, coder_ID, general_code)
ar_general_2.1 <- pivot_wider(ar_general_2.1, id_cols = coder_ID,
names_from = content_ID, values_from = general_code)
ar_general_2.1 <- select(ar_general_2.1, -coder_ID)

#Arabic coders specific perception of OI input
ar_specific_2.1 <- select(irr_2.1ar, content_ID, coder_ID,
specific_code)
ar_specific_2.1 <- pivot_wider(ar_specific_2.1, id_cols = coder_ID,
names_from = content_ID, values_from = specific_code)
ar_specific_2.1 <- select(ar_specific_2.1, -coder_ID)

#Arabic coders specific belief of OI importance
ar_specific_2.2 <- select(irr_2.2ar, content_ID, coder_ID,
specific_code)
ar_specific_2.2 <- pivot_wider(ar_specific_2.2, id_cols = coder_ID,
names_from = content_ID, values_from = specific_code)
ar_specific_2.2 <- select(ar_specific_2.2, -coder_ID)
```

In order to calculate Krippendorff's Alpha for each variable, the data frames are transformed into matrices and reliability is calculated using the "kripp.alpha" function of the "irr" package. The output for the intercoder reliability for the categories within each variable is reported in table XIII.2.

```
en_general_2.1 <- as.matrix(en_general_2.1)
kripp.alpha(en_general_2.1, method = "nominal")

en_specific_2.1 <- as.matrix(en_specific_2.1)
kripp.alpha(en_specific_2.1, method = "nominal")

en_specific_2.2 <- as.matrix(en_specific_2.2)
kripp.alpha(en_specific_2.2, method = "nominal")
```

Table XIII.2: Krippendorff's alpha (α) assessing intercoder reliability across OI influence and importance cateogries, by L1

Category agreement (α)	**L1 English (*n*=3)**	**L1 Arabic (*n*=3)**
Q6 general	1 (n=27)	1 (n=26)
Q6 specific	0.71 (n=27)	0.85 (n=26)
Q8 specific	0.76 (n=25)	0.76 (n=25)

Appendix XIV: Coded questionnaire responses for Q6

Q 6: Did it make a difference seeing the written words? In what way? **عندما تتعلم كلمات جديدة هل أحدث ذلك فرقا في رؤية الكلمات المكتوبة؟ إلي أي مدي؟**		
Group	**ID**	**Response [with tags]**
Arabic	5000814	فرق كبير*[EXCL - INSUFFICIENT]*
Arabic	5029581	No [DOESN'T HELP]
Arabic	5082295	نعم ، لمدى كثير جدا بالمكتوبه استطيع معرفه النطق الصحيح [HELPS - CLARIFY WHAT I HEARD]
Arabic	5229722	As some of the words sound similar, seeing it wtitten helped me [HELPS] distinguish between them when hearing them so I would say it helped a lot in that case [DIFFERENTIATE BETWEEN SIMILAR SOUNDS/WORDS], it also helped with other cases as linking the image and the sound with a written word made it easier to remember [ASSOCIATE WORD AND IMAGE].
Arabic	5336078	yes, alot *[EXCL - INSUFFICIENT]*
Arabic	5336722	Maybe it has made it easier a little bit, but I relied more on memorizing the shape and the word heard [DOESN'T HELP - BETTER TO IGNORE/ FOCUS ON SOUNDS].
Arabic	5359092	فرق كبير جدا *[EXCL - INSUFFICIENT]*
Arabic	5338565	it is help some time to remember [HELPS - REMEMBERING]
Arabic	5341799	yes, definitely *[EXCL - INSUFFICIENT]*
Arabic	5378097	تقريبا *[EXCL - INSUFFICIENT]*
Arabic	5442543	بعض الكلمات لم ألاحظ أنها تبدأ بصوتين متشابهين مثلا v and f وبما أن هذه الأصوات غائبة في لغتي في الغالب لن أنتبه لها عند سماعها فقط فالإملاء يساعد في هذه الحالة [HELPS - DIFFERENTIATE SIMILAR SOUNDS/WORDS]
Arabic	5500029	Far *[EXCL - INSUFFICIENT]*
Arabic	5501560	wide range *[EXCL - INSUFFICIENT]*
Arabic	5730442	لا يوجد فرق كبير[DOESN'T HELP]
Arabic	5607775	i don't now *[EXCL - INSUFFICIENT]*
Arabic	5359752	نعم،كتابة الكلمة باللغة الانجليزية تجعل تعلمها اسهل [HELPS - LEARNING]

Arabic	5089629	#N/A
Arabic	5622524	فرق بنسبة 75% عن الكلمات الغير مكتوبه *[EXCL - INSUFFICIENT]*
Arabic	5439128	إلى حد ما *[EXCL - INSUFFICIENT]*
Arabic	5059373	yes it did because i could remember the words easier [HELPS - REMEMBERING]
Arabic	5029587	for English written forms are helpful [HELPS] as English is not a consistent language (written forms to spoken forms) [MAP SOUNDS AND LETTERS]
Arabic	5037046	الكلمات المكتوبة مصدر تأكيد للكلمة المسموعة [HELPS - CLARIFY WHAT I HEARD]
Arabic	5553283	Yas *[EXCL - INSUFFICIENT]*
Arabic	5015277	yes ,i prefer the word written while listening to it [HELPS], that's how i memorize the word twice written and spoken [CONNECT AUDIO AND VISUAL INPUT]
Arabic	5462535	to the remember of the words [HELPS - REMEMBERING]
Arabic	5011969	نعم أحدث فرقا إلي مدى كبير *[EXCL - INSUFFICIENT]*
Arabic	5499711	3 *[EXCL - INSUFFICIENT]*
Arabic	5365353	No [DOESN’T HELP]
Arabic	5500050	just a little *[EXCL - INSUFFICIENT]*
Arabic	5499774	Yes, How much it helps to learn. [HELPS - LEARNING]
Arabic	5499647	نعم، الكلمات المكتوبة كانت اصعب [DOESN'T HELP]
Arabic	5450539	في الغالب لا يحدث فرق بل ربما يشتت قليلا النظر الى الكلمة المكتوبة والصورة والاستماع [DOESN'T HELP - DISTRACTING (FROM SOUNDS)]
Arabic	5452186	helped me remember the words as I am a visual person [HELPS - REMEMBERING]
Arabic	5384016	متوسط *[EXCL - INSUFFICIENT]*
Arabic	5370467	كانت الكلمات المكتوبة باللغة الإنجليزية أدعى لتذكرها والربط بينها وبين الصورة [HELPS - ASSOCIATE WORD AND IMAGE]
Arabic	5370466	لا، أركز فقط على الصوت [DOESN'T HELP - BETTER TO FOCUS ON SOUNDS]

Arabic	5369858	إلى نسبة 50% *[EXCL - INSUFFICIENT]*
Arabic	5370516	رؤية الكلمة الإنجليزية يساعدني [HELPS] أحيانا في استيعاب المسموع وخاصة أن الكلمات الإنجليزية تطابق حروفها الصوت المسموع. قد تؤدي الكلمات العربية الدور نفسه لو كانت الكلمات العربية مشكلة بالحركات [MAP SOUNDS AND LETTERS]
Arabic	5359087	لا [DOESN'T HELP]
Arabic	5359091	اذا كانت الكلمة مكتوبة بالانجليزية اسهل [HELPS - LEARNING]
Arabic	5337282	yes it takes time to remember the words and make a connection between each shape and its related words. [NEGATIVE - DISTRACTING (FROM SOUNDS)]
Arabic	5357676	نعم فقد ساعدني [HELPS] في تذكر شكل الكلمة كثيرا سواء الكلمة بالعربية او الانجليزية فأنا أتخيلها واتذكر لو حتى حرفا واحدا [VISUALISE WORDS]
Arabic	5342157	استيعاب لكثر [HELPS - LEARNING]
Arabic	5341797	نعم، لأنني أربط بين شكل الكلمة والصورة [HELPS - ASSOCIATE WORDS AND PICTURE]
Arabic	5339892	نعم، أصبحت الفروق واضحة بين F و V. [HELPS - DIFFERENTIATE SIMILAR SOUNDS/ WORDS] كما كانت القراءة والاستماع في آن وحدا وسيلة للتأكيد على حفظ الكلمة. [CONNECT AUDIO AND VISUAL INPUT]
Arabic	5335791	رؤية الكلمات العربية سهل تذكرها بالنسبة لي اكثر من الانجليزية [HELPS - REMEMBERING]
Arabic	5339069	Yes, I feel that I can remember the written words easier [HELPS -REMEMBERING]
Arabic	5337115	1 *[EXCL - INSUFFICIENT]*
Arabic	5341802	نعم *[EXCL - INSUFFICIENT]*
Arabic	5336140	sometime i get distracted by the written word as it does not make me focus as much on the picture. [DOESN'T HELP - DISTRACTING (FROM SOUNDS)]
Arabic	5338152	نعم *[EXCL - INSUFFICIENT]*
Arabic	5338811	نعم *[EXCL - INSUFFICIENT]*
Arabic	5337579	نعم أحدث فرقا كبيرا *[EXCL - INSUFFICIENT]*
Arabic	5338948	جعل تذكرها اسهل [HELPS - REMEMBERING]

Arabic	5337166	الفرق بدى واضحًا في الكلمات المتشابهه مثل Fan Van [HELPS DIFFERENTIATE SIMILAR SOUNDS/ WORDS]
Arabic	5337506	اسهل في القراءه [HELPS - LEARNING]
Arabic	5336602	ربط الصور مع الكلمات يجعل عمليه التذكر اسهل [HELPS - ASSOCIATE WORD AND IMAGE]
Arabic	5337014	I don't know *[EXCL - INSUFFICIENT]*
Arabic	5336601	writing words in Arabic made it somehow challenging to learn them. [DOESN'T HELP]
Arabic	5335700	Yes, it created a mental image which made choosing easier. [HELPS - VISUALISE WORDS]
Arabic	5331395	Yes *[EXCL - INSUFFICIENT]*
Arabic	5299874	0.7 *[EXCL - INSUFFICIENT]*
Arabic	5298847	sometimes it helps and sometimes it doesn't *[EXCL - INSUFFICIENT]*
Arabic	5212730	نعم *[EXCL - INSUFFICIENT]*
Arabic	5247103	yes. It was easier to remember the words. [HELPS - REMEMBERING]
Arabic	5246620	not much [DOESN'T HELP]
Arabic	5155297	Sometime *[EXCL - INSUFFICIENT]*
Arabic	5196532	connection between the letters and pronunciation might help to remeber. [HELPS - MAP SOUNDS AND LETTERS]]I am not sure if I understood the question.
Arabic	5201553	You can learn better when you see the words rather when you hear them. [HELPS - LEARNING]
Arabic	5212731	الاستماع والقراءة معا تجعل من التعلم أسهل [HELPS - CONNECT AUDIO AND VISUAL INPUT]
Arabic	5182923	Yes *[EXCL - INSUFFICIENT]*
Arabic	5201938	to remember the word [HELPS - REMEMBERING]
Arabic	5182922	نعم- ساعدتني كثيرا الكلمات المكتوبة [HELPS - LEARNING]
Arabic	5189512	It made it easier to remember the words [HELPS - REMEMBERING]

Arabic	5155797	No *[EXCL - INSUFFICIENT]*
Arabic	5089630	yes, it makes it easier to remembr the word because you saw its spelling [HELPS - REMEMBERING]
Arabic	5097676	يجعل التذكر اسرع [HELPS - REMEMBERING]
Arabic	5029580	Yes, it makes learning easier [HELPS - LEARNING]
Arabic	5082294	لا [DOESN'T HELP]
Arabic	5018214	0.4 *[EXCL - INSUFFICIENT]*
Arabic	5003021	الى مدى صغير *[EXCL - INSUFFICIENT]*
Arabic	5027737	yes it did. Seeing the written words made it easier to remember. [HELPS - REMEMBERING]
Arabic	5010776	Notba big difference [DOESN'T HELP]
Arabic	5047738	نوعا ما *[EXCL - INSUFFICIENT]*
Arabic	5038591	التعلم اسرع بوجود الكلمات المكتوبة [HELPS - LEARNING]
Arabic	5044345	بنسبة كبيرة *[EXCL - INSUFFICIENT]*
Arabic	5037277	much better [HELPS - LEARNING]
Arabic	5029573	كتابه الكلمات يسهل الحفظ [HELPS - REMEMBERING]
Arabic	5010775	no difference. hearing words make it easier than reading them, especially in Arabic [DOESN'T HELP - BETTER TO IGNORE/ FOCUS ON SOUND]
Arabic	5034149	it's very helpful when I see the word written [HELPS - LEARNING]
Arabic	5016049	الى مدى كبير *[EXCL - INSUFFICIENT]*
Arabic	5031281	watching the written words make learning easier [HELPS] as I can link the image with the written word [ASSOCIATE WORD AND IMAGE]
Arabic	5031282	نعم .. إلى مدى كبير *[EXCL - INSUFFICIENT]*
Arabic	5029568	نعم إلى مدى مهم في تنشيط الذاكرة والحفظ [HELPS - REMEMBERING]
Arabic	5029577	0.5 *[EXCL - INSUFFICIENT]*
Arabic	5029572	لا [DOESN'T HELP]

Arabic	5029579	No [DOESN’T HELP]
Arabic	5027663	نعم الكتابة تساعد على الربط مع الصوت [HELPS - MAP SOUNDS AND LETTERS]
Arabic	5027661	yes made a difference , but not that much *[EXCL - INSUFFICIENT]*
Arabic	5027570	i can remember the word spelling, so each time i see the picture i remember the word, but this didn't happen to me in all the pictures only in some of them. [HELPS - REMEMBERING]
Arabic	5003012	نعم، الى مدى متوسط *[EXCL - INSUFFICIENT]*
Arabic	5003019	not too much [DOESN’T HELP]
Arabic	5010884	الى حد كبير *[EXCL - INSUFFICIENT]*
Arabic	5003023	لا لا يوجد فرق [DOESN'T HELP]
Arabic	5010779	كبير *[EXCL - INSUFFICIENT]*
Arabic	5003015	Yes it did just to distinguish between the letters that are close in pronunciation [HELPS - DIFFERENTIATE SIMILAR SOUNDS/WORDS]
Arabic	5000809	الكلكات المكتةبة تساعد [HELPS] على التذكر بقوة [REMEMBERING]
Arabic	5000812	Huge *[EXCL - INSUFFICIENT]*
Arabic	5000759	كبير *[EXCL - INSUFFICIENT]*
Arabic	5000407	نعم *[EXCL - INSUFFICIENT]*
Arabic	4988182	Yes *[EXCL - INSUFFICIENT]*
Arabic	4986648	differentiate betweet the V and F [HELPS - DIFFERENTIATE SIMILAR SOUNDS/ WORDS]
Arabic	4986649	written words are much easier to remember and to associate with the right item [HELPS - ASSOCIATE WORD AND IMAGE]
Arabic	4986647	نعم. انا انسانة بصرية [HELPS - LEARNING]

Q 6: Did it make a difference seeing the written words? In what way? **عندما تتعلم كلمات جديدة هل أحدث ذلك فرقا في رؤية الكلمات المكتوبة؟ إلي أي مدي؟**		
Group	**ID**	**Response [with tags]**

English	5304273	yes - easier to help remember when in english [HELPS - REMEMBERING]
English	5305661	Easier to visualise [HELPS - VISUALISE WORDS]
English	5308502	i think it aided the distinction between similar sounding words [HELPS - DIFFERENTIATE SIMILAR SOUNDS/ WORDS]
English	5308657	Yes, in English because it helped me to clarify if I had heard the spoken pronunciation correctly. I was often surprised at the spelling (e.g. "nadus" spelt nad-I-s). [HELPS - CLARIFY WHAT I HEARD]
English	5308960	In english, enabled me to spell it out in my head [HELPS - MAP SOUNDS AND LETTERS]
English	5312160	It makes a difference to me in general because I can see the sounds I am unable to pick up by just listening. [HELPS - CLARIFY WHAT I HEARD]
English	5312301	Yes - dual coding [HELPS - CONNECT AUDIO AND VISUAL INPUT]
English	5312548	I found English written words slightly helpful [HELPS], while Arabic written words were more of a distraction. [DOESN'T HELP - DISTRACTING (FROM SOUNDS)]
English	5315514	Much easier to remember [HELPS - REMEMBERING]
English	5322227	It was easier to distinguish them [HELPS - DIFFERENTIATE SIMILAR SOUNDS/ WORDS]
English	5323642	I think seeing the written words alongside hearing the new words was helpful [HELPS]. This is becuase some of the new words sounded very similar to each other... letters that have a similar English pronunciation such as 'v & f', or, 'm & n', when in a similar word are slighlty tricker to distinguish from one another on hearing them for the first time [DIFFERENTIATE SIMILAR SOUNDS/ WORDS]. Even first letters that are less likely to become muddled, when in a similar word structure, (like Nadis and Famis) becomes slightly confusing. Seeing the word written out, helped assure me of what the spoken was [CLARIFY WHAT I HEARD].
English	5327303	I was concentrating almost purely on the sound of words [DOESN'T HELP - BETTER TO IGNORE/ FOCUS ON SOUNDS]
English	5328002	Yes, easier to associate the word with the object when you can see the spelling [HELPS - ASSOCIATE WORD AND IMAGE]

English	5330132	easier to remember [HELPS - REMEMBERING]
English	5332522	something else to help me remember! [HELPS - REMEMBERING]
English	5332693	In english yes, as could actively read the word while it was being said. [HELPS - MAP SOUNDS AND LETTERS]
English	5333575	It helped [HELPS] to pair the visual with the sounds in my head [MAP SOUNDS AND LETTERS] as I would try to say them if they were written in English [REHEARSE WORDS].
English	5335169	As many of the words sounded the same, seeing them written helped [HELPS - DIFFERENTIATE SIMILAR SOUNDS/ WORDS]. Although still hard to remember which is which after the 2 object game
English	5336201	it helps to picture the word when looking at the picture so they become like a pair in my mind [HELPS - ASSOCIATE WORD AND IMAGE]
English	5337460	It made it easier to distinguish between different words by associating the picture with how the word was written, especially when two words sounded very similar. [HELPS - DIFFERENTIATE SIMILAR SOUNDS/ WORDS]
English	5338550	i felt it was easier to remember the picture and the word together when they flashed up on scren together - like trying to take a mental screenshot [HELPS - ASSOCIATE WORD AND IMAGE]
English	5338600	Seeing english written words helped me differitiate between the different similar words,ie, helped notice if it ended in T rather S [HELPS - DIFFERENTIATE SIMILAR SOUNDS/ WORDS]
English	5339584	yes, helps to understand and triggers visual memory [HELPS - REMEMBERING]
English	5339906	helped figure out the subtle differences in pronunciations [HELPS - DIFFERENTIATE SIMILAR SOUNDS/ WORDS]
English	5340413	Seeing the letters m/n and v/f helped me know these words better for when the 'test'. [HELPS - FOCUS ON FIRST LETTER]
English	5341833	Helped with initial consonent (e.g. v/f, m/n) [HELPS - DIFFERENTIATE SIMILAR SOUNDS/ WORDS]
English	5341870	It made a difference in English because my eyes were naturally drawn to the English words. To be honest, I ignored the Arabic script and just focused on the pictures [DOESN’T HELP - BETTER

		TO IGNORE/ FOCUS ON SOUND].
English	5343885	When they were written in English it made it easier to remember for me [HELPS - REMEMBERING]
English	5344305	Not for me, I read them when I noticed them but it didn't seem to help [DOESN'T HELP], I tried to visually attach the sound to the image or something that might help me remember.... Like the 'famus' (sp!) sounded like 'female' and the item looked female! [BETTER TO IGNORE/ FOCUS ON SOUNDS]
English	5353998	No [DOESN'T HELP]
English	5354235	no as i didnt understand tge arabic writing [DOESN'T HELP - UNFAMILIAR]
English	5356536	no difference [DOESN'T HELP]
English	5356537	much easier in english, and seeing the spelling helped in remembering [HELPS - REMEMBERING]
English	5357670	It was visually associative, whether in Arabic or English [HELPS - CONNECT AUDIO AND VISUAL INPUT]. However, sometimes with the English words, I concentrated more on the visual aspect of the word rather than the sound [DOESN'T HELP - DISTRACTING (FROM SOUNDS)]
English	5359677	No [DOESN'T HELP]
English	5366758	It made it easier to comprehend [HELPS - CLARIFY WHAT I HEARD] the sound, and recall it [REMEMBERING]
English	5366866	Yes *[EXCL - INSUFFICIENT]*
English	5374409	I could visualise what I was hearing later [HELPS - VISUALISE WORDS]
English	5374964	yes, i think it helped me to remember [HELPS - REMEMBERING]
English	5380819	No [DOESN'T HELP]
English	5384030	the English words helped my brain link the images to the sound better [HELPS - ASSOCIATE WORD AND IMAGE]
English	5384032	some words were very similar so it made it easier to see the spelling [HELPS - DIFFERENTIATE SIMILAR SOUNDS/ WORDS]
English	5384657	None [DOESN'T HELP]

English	5409941	I didn't notice the written words as much as hearing them [DOESN'T HELP]. The written words were most a distraction [DISTRACTING (FROM SOUNDS)], and as this test appeared to be timed, it seemed ineffecient to add trying to learn the written version of the word in addition to its sound, so I mostly ignored the written words [BETTER TO IGNORE/ FOCUS ON SOUNDS].
English	5417849	Yes, helped to know which letter each word started with. [HELPS - FOCUS ON FIRST LETTER]
English	5418540	Seeing the word made it easier to understand what the word was, because even if you misheard the word, the spelling was there so reinforce what the word was [HELPS - CLARIFY WHAT I HEARD]. It also meant you heard the word and saw the word so you had two chances to learn the word [CONNECT AUDIO AND VISUAL INPUT].
English	5434242	The change in language distracted from the image and sound of the word *[EXCL - TASK PROBLEM]*
English	5476814	Highlighted the v/f difference! N/M seemed more obvious out loud. [HELPS - DIFFERENTIATE SIMILAR SOUNDS/ WORDS]
English	5499381	no difference [DOESN'T HELP] did not look at the words trying to remember what the image looks like [BETTER TO IGNORE/ FOCUS ON SOUNDS]
English	5499420	Yes, seeing them helped me rehearse the word in my head [HELPS - REHEARSE]. It also confirmed the sounds that the speaker was making (e.g. v versus f) [DIFFERENTIATE SIMILAR SOUNDS/ WORDS]
English	5499421	I don't think so particularly. Arabic text was of no help and I don't think the English text made a noticable difference. [DOESN'T HELP]
English	5502013	seeing the initial letter helped. [HELPS - FOCUS ON FIRST LETTER]
English	5503474	Yes, it helped me to hear the separation between [m] and [n] or [v] and [f] [HELPS - DIFFERENTIATE SIMILAR SOUNDS/ WORDS]
English	5510526	With the English ones, it made it slightly easier to remember them I think as I sometimes I would match the first syllable with a letter in my memory [HELPS - FOCUS ON FIRST LETTER]. It made no difference whatsoever with the arabic words because I cannot differentiate between them anyway [DOESN'T HELP - UNFAMILIAR].
English	5521508	I preferred learning from sound alone, as English spelling didn't

		reflect pronunciation [DOESN'T HELP - BETTER TO IGNORE/ FOCUS ON SOUNDS].
English	5557001	Seeing English spelling helped to make connections between what I was hearing and what I was seeing [HELPS - CONNECT AUDIO AND VISUAL INPUT]
English	5586643	found it easier in understanding spoken word which in turn helped identify object [HELPS - CLARIFY WHAT I HEARD]
English	5586843	The writing didn't help or hinder in any way - I focused more on hearing them out loud and trying to match the sound to the colours and shapes. [DOESN'T HELP - BETTER TO IGNORE/ FOCUS ON SOUNDS]
English	5587514	No [DOESN'T HELP]
English	5588274	i dont remember seeing any written words! feel a bit silly now…[DOESN'T HELP - BETTER TO IGNORE/ FOCUS ON SOUNDS]
English	5592006	it made a small difference, i was able to distinguish more easily the subtle differences between similar words [HELPS - DIFFERENTIATE SIMILAR SOUNDS/ WORDS]
English	5595978	The English words I understood so it didn't make too much difference [DOESN'T HELP], but the Arabic words put with the picture made it so that I felt there was more to take in, so distracted from learning the sound and I was trying to correlate th sound to the shape of the word?!! [DOESN'T HELP - DISTRACTING]
English	5602966	It helped consolidate my learning, it acted as another source for me to remember and help me learn the words [HELPS - CONNECT AUDIO AND VISUAL INPUT]. It also helped clarify sounds which might be similar, e.g. was it an f or v that was said [DIFFERENTIATE SIMILAR SOUNDS/ WORDS].
English	5615580	yes, it made it easier as you can match spelling to sound in some cases [HELPS - MAP SOUNDS AND LETTERS]
English	5620967	I could read the word out loud to myself [HELPS - REHEARSE WORDS]
English	5621683	No. I was only listening [DOESN'T HELP - BETTER TO IGNORE/ FOCUS ON SOUNDS]
English	5627982	No [DOESN'T HELP]

English	5627983	not really [DOESN'T HELP]
English	5629029	Linked the spelling to the object, eg nasit - to me looked like a nest i could sit on [HELPS - ASSOCIATE WORD AND IMAGE]
English	5637584	no difference [DOESN'T HELP]
English	5642980	Yes, I was able to visualise the word with the object [HELPS - ASSOCIATE WORD AND IMAGE]
English	5649032	It helped to see the English versions written out to match sounds of pronunciation with what I was hearing [HELPS - MAP SOUNDS AND LETTERS]. But without words I focused more on the sounds and being able to differentiate eg similar prefixes.
English	5655323	No [DOESN'T HELP]
English	5659195	yes when work were in English, seeing the spelling made it easier to remember e.g what letter it started with [HELPS - FOCUS ON FIRST LETTER]
English	5659326	yes, it was easier to learn the similar sounding words when I was able to see the starting letter. [HELPS - DIFFERENTIATE SIMILAR SOUNDS/ WORDS]
English	5663887	In English it helped because it gave me something else to remember and map the sound onto [HELPS - MAP SOUNDS AND LETTERS]. In Arabic, I just ignored it so it made no difference [DOESN'T HELP - BETTER TO IGNORE/ FOCUS ON SOUNDS].
English	5664502	For English, being able to distinguish minimal pairs, e.g., m/n and f/v [HELPS - DIFFERENTIATE SIMILAR SOUNDS/ WORDS]
English	5666899	Seeing written English helped differentiate the first letter on words that sounded similar [HELPS - DIFFERENTIATE SIMILAR SOUNDS/ WORDS]
English	5667662	No [DOESN'T HELP]
English	5670727	no, i found myself listening to the pronunciation and relating the sounds to the pictures rather than looking at the spellings [DOESN'T HELP - BETTER TO IGNORE/FOCUS ON SOUNDS]
English	5671003	I wasn't really paying attention to the written words at first so I thought it made no difference, but on some level I must have been paying attention because when they were no longer included, it became a little more difficult to hear what letters were being said [HELPS - CLARIFY WHAT I HEARD].

English	5672324	Yes - it helped me get a picture of the word in my head to remember [HELPS - VISUALISE WORDS]
English	5674519	Yes *[EXCL - INSUFFICIENT]*
English	5676564	Yes I was able to be sure of whether the speaker said m/n or v/f as they can sound quite similar [HELPS - DIFFERENTIATE SIMILAR SOUNDS/ WORDS]. i was also able to remember the spelling and repeat. hearing the word aloud from myself helped me [REHEARSE WORDS]
English	5690497	Yes it is easier for me to remember words if I see them written. [HELPS - REMEMBERING]
English	5691858	Yes. It was easier to remember the object described. [HELPS REMEMBERING]
English	5692078	Helpful when in English [HELPS - LEARNING]
English	5692100	Yes. I had more to associate the sound with the picture [HELPS - ASSOCIATE WORD AND IMAGE]
English	5696781	I don't think so [DOESN'T HELP]
English	5704997	Yes, i feel it triggered my memory [HELPS - REMEMBERING]
English	5706490	yes. accentuates the difference between f / v sounds, m / n sounds [HELPS - DIFFERENTIATE SIMILAR SOUNDS/ WORDS]
English	5706491	In English it helped because I could begin to match them together when I saw them repeatedly over time. [HELPS - CONNECT AUDIO AND VISUAL INPUT]
English	5719579	It confirmed what I thought I was hearing [HELPS - CLARIFY WHAT I HEARD]
English	5732647	Helped visual memory linked to the image [HELPS - ASSOCIATE WORD AND IMAGE]
English	5750023	I'm not 100% sure. I don't think it made much of a difference though [DOESN'T HELP]
English	5750994	Provided context to the pronunciation and provided more objective concepts in my head of what the thing was than audio alone [HELPS - CLARIFY WHAT I HEARD]
English	5751031	Because there weren't consistently written words in a language I read [DOESN'T HELP - UNFAMILIAR], I decided early on not to focus on spelling/written language, only on the auditory. [BETTER

		TO IGNORE/ FOCUS ON SOUNDS]
English	5755297	Yes, I am a visual learner, so being able to connect the English spelling of the words with the images helped me differentiate the images with similar sounds [HELPS - ASSOCIATE WORD AND IMAGE]
English	5755298	It didn't make too much of a difference since I could typically hear the difference between similar sounds (f/v), although seeing the English written did clarify the endings of the word for me [HELPS - CLARIFY WHAT I HEARD]. I didn't pay much attention to the Arabic written words [DOESN'T HELP - BETTER TO IGNORE/ FOCUS ON SOUNDS]]and seeing no words was fine.
English	5755299	not particularly. i memorized thd phonetic content eithout paying too much attention to orthography [DOESN'T HELP - BETTER TO IGNORE/ FOCUS ON SOUNDS]
English	5755300	When I noticed the words written in English, it didn't help as much as I expected; I mainly went through associating the sounds with the image [DOESN'T HELP - BETTER TO IGNORE/ FOCUS ON SOUNDS]
English	5755301	the writing helped to clarify which initial sound the word had (v or f, m or n, etc) [HELPS - DIFFERENTIATE SIMILAR SOUNDS/ WORDS]
English	5756084	It was easier to identify the sound I heard was correct [HELPS - CLARIFY WHAT I HEARD]
English	5756085	English transliterations helped bond sound and sense [HELPS - ASSOCIATE WORD AND IMAGE]
English	5758767	easier when they were written in English - I think it was easier to remember the first letter [HELPS - FOCUS ON FIRST LETTER]
English	5759398	easier to distinguish start letter [HELPS - FOCUS ON FIRST LETTER]
English	5759735	yes, it helped with pronouncing [HELPS - REHEARSE WORDS]
English	5762777	it was easier to differentiate the consonant sounds after seeing the spelling in english [HELPS - DIFFERENTIATE SIMILAR SOUNDS/ WORDS]
English	5765135	it helped put some letters to a sound, this made it easier because it was remembering more distinct qualities related to the sound that was played [HELPS - MAP SOUNDS TO LETTERS]

English	5765136	The words accompanied by the English spelling were easier to remember because it was a familiar way of learning new words in general. I find it easier to remember words by their spelling rather than by an image [HELPS - REMEMBERING]. The Arabic spelling didn't make any difference because I do not know how to read it [DOESN'T HELP - UNFAMILIAR], the only thing I did was trying to see a similarity between the shape of the words and the sound of them.
English	5780967	easier to recall [HELPS - REMEMBERING]
English	5831108	was able to picture which picture i was thinking with the spelling when doing the test [HELPS - ASSOCIATE WORD AND IMAGE]
English	5884960	No [DOESN'T HELP]
English	5909329	reinforces the sound of the word [HELPS - CLARIFY WHAT I HEARD]
English	5909763	Yes - helpful to gain a sense of pronunciation and structure [HELPS - CLARIFY WHAT I HEARD]
English	5914367	No [DOESN'T HELP]
English	5998960	English text helped with visualising the word [HELPS - VISUALISE WORDS]
English	6029761	Maybe helped for remembering the first letter of each word [HELPS - FOCUS ON FIRST LETTER]

Appendix XV: Coded questionnaire responses for Q7 and Q8

Q7: In general, when learning new words do you think it is important to see the spelling the first time you hear it?

بشكل عام، عندما تتعلم كلمات جديدة هل من المهم عندك أن ترى املاء الكلمة عند الاستماع لها لأول مرة ؟

Q8: Why?

لماذا؟

Group	ID	Response [with tags]	
Arabic	5000814	yes	i may hear wrong [CORRECT COMPREHENSION]
Arabic	5029581	yes	to pronunce it correctly [CORRECT PRODUCTION]
Arabic	5082295	no	ليس مهم في هذا الوقت *[EXCL - INSUFFICIENT]*
Arabic	5229722	yes	This way I got more information which means more links to the word to easily recall it. [VISUAL ASSOCIATION]
Arabic	5336078	yes	it helps me remember the word [REMEMBERING]
Arabic	5336722	yes	To realize how it's pronounced later when I see the word. [CORRECT PRODUCTION]
Arabic	5359092	yes	افهم الكلمة اكثر [CORRECT COMPREHENSION]
Arabic	5338565	yes	To say it correct [CORRECT PRODUCTION]
Arabic	5341799	yes	to mach what I'm listening to [VISUAL ASSOCIATION]
Arabic	5378097	yes	لكي يسهم البصر في حفظ الكلمة [MEMORISATION]
Arabic	5442543	yes	لصعوبة التفريق أحيانا بين الأصوات [CORRECT COMPREHENSION]
Arabic	5500029	yes	to be familiar with that word [CLARIFYING AND CONSOLIDATING NEW LANGUAGE]
Arabic	5501560	yes	to memorize it in better way [MEMORISATION]
Arabic	5730442	yes	لأستطيع تمييزها عن الكلمات المتشابهة لأن الاستماع للغة الانكليزية غير واضح [CORRECT COMPREHENSION]
Arabic	5607775	no	because letters silent [SOUND IS MORE IMPORTANT/ BENEFICIAL]
Arabic	5359752	yes	حتى اربط بين املاء الكلمة وصوتها [DECODING SOUNDS AND LETTERS]
Arabic	5089629	#N/A	#N/A

Arabic	5622524	yes	كي اعرف الاسم الصحيح والنطق [CORRECT PRODUCTION]
Arabic	5439128	yes	حتى يكون هناك ربط بين الصوت ورسم الكلمة [VISUAL ASSOCIATION]
Arabic	5059373	yes	i am a visual learner [LEARNING STYLE] and seeing the words written help me to memorize better [MEMORISATION]
Arabic	5029587	yes	for recall support [REMEMBERING]
Arabic	5037046	yes	هو بمثابة تثبيت للكلمة، [CLARIFYING AND CONSOLIDATING NEW LANGUAGE]
Arabic	5553283	yes	للربط بين الصورة والصوت والكتابة *[EXCL - ABOUT THE EXPERIMENT]*
Arabic	5015277	yes	to memorize it better [MEMORISATION]
Arabic	5462535	yes	To understand the word more [CORRECT COMPREHENSION] and know how to properly use it [CORRECT PRODUCTION]
Arabic	5011969	yes	لتسهيل نطقها وترسيخه [CORRECT PRODUCTION]
Arabic	5499711	yes	It is better *[EXCL - INSUFFICIENT]*
Arabic	5365353	no	sound moreimportant [SOUND IS MORE IMPORTANT/ BENEFICIAL]
Arabic	5500050	yes	to be familiar with that word [CLARIFYING AND CONSOLIDATING NEW LANGUAGE]
Arabic	5499774	yes	That help me to learn quickly [SPEED OF LEARNING]
Arabic	5499647	no	لكي لا اتشتت بالتركيز على املاء تلك الكلمة [IT'S TOO MUCH TO PROCESS]
Arabic	5450539	no	لانه يصعب علي المهمة في حفظ الاملاء للاحرف وحفظ الكلمة المستمعة [IT'S TOO MUCH TO PROCESS]
Arabic	5452186	yes	I am a visual person [LEARNING STYLE]. it helps me remember [REMEMBERING]
Arabic	5384016	yes	لتثبت في مخيلتي [VISUALISATION]
Arabic	5370467	yes	لأنه أدعى لتذكرها [REMEMBERING]
Arabic	5370466	yes	لأعرف نطق الكلمة [CORRECT PRODUCTION]

Arabic	5369858	yes	لتمييزها عند قرائتها [CORRECT COMPREHENSION]
Arabic	5370516	yes	كما أشرت في الإجابة على سؤال سابق،، أولا: الكلمات العربية تحتاج إلى حركات التشكيل لتؤدي نفس الغرض الذي تؤديه الكلمات الإنجليزية للصوت نفسه، وثانية في اللغة الإنجليزية أصوات غير متوفرة في حروف العربية، فمثلا الأحرف (f and v) العربية لا تفرق بين صوتي [DECODING SOUNDS AND LETTERS]
Arabic	5359087	yes	أحيانًا أربط بين الشكل والصوت أو المعنى [VISUAL ASSOCIATION]
Arabic	5359091	yes	ترتبط الصورة بشكل املاء الكلمة *[EXCL - ABOUT THE EXPERIMENT]*
Arabic	5337282	yes	its help me to draw a vision on my mind and link between the word and the shape [VISUALISATION]
Arabic	5357676	yes	يساعدني جدا في تذكرها [REMEMBERING] وربطها بكلمة في لغتي الام ولو برابط مضحك خاص بي انا .. لا اخبره لأحد [VISUAL ASSOCIATION]
Arabic	5342157	yes	حتى استطيع الربط [VISUAL ASSOCIATION] ويسهل المفظ [MEMORISATION] والتأكد من الاملاء للكلمة [CLARIFYING AND CONSOLIDATING NEW LANGUAGE]
Arabic	5341797	yes	حتى أعرف النطق الصحيح، لأنني أحيانا لا أفرق بين f و v .فرؤية الكلمات مكتوبة يساعد على التعلم *[EXCL - ABOUT THE EXPERIMENT]*
Arabic	5339892	yes	حتى أفرق بينها وبين كلمة أخرى شبيهة ، وحتى أتعرف عليها في سياق القراءة[CORRECT COMPREHENSION]، ولأتمكن من استخدامها في التحدث والكتابة [CORRECT PRODUCTION].
Arabic	5335791	yes	أحياناً أتذكر شكل الكلمة فيساعدني على تذكر الكلمة نفسها مثلا أثناء الاختبارات المدرسية [VISUALISATION]
Arabic	5339069	yes	because it helps in remembering the words, therefore the next time I want to use this word, I visualize how it is written in my brain [VISUALISATION] and then use it [CORRECT PRODUCTION].
Arabic	5337115	yes	لاحفظها بسرعة [SPEED OF LEARNING]
Arabic	5341802	yes	ربط بين الصوت والشكل [VISUAL ASSOCIATION]
Arabic	5336140	no	beacuse sometimes writing does not refelct how we speak the language. there may be little correspondnce between the writing systme and the sound of a language. [SPELLING DOESN’T REFLECT PRONUNCIATION]
Arabic	5338152	yes	to know how it is written [CORRECT PRODUCTION]
Arabic	5338811	yes	حتى ترتكز في المخ [CLARIFYING AND CONSOLIDATING NEW LANGUAGE]

Arabic	5337579	yes	لافهم اكثر [CORRECT COMPREHENSION]
Arabic	5338948	yes	لا اعلم *[EXCL - INSUFFICIENT]*
Arabic	5337166	yes	يسهل تذكر شكل الحروف [VISUALISATION]
Arabic	5337506	yes	لاتذكر شكلها في ذهني [VISUALISATION]
Arabic	5336602	yes	#N/A
Arabic	5337014	yes	I can recognise it and remember it faster [SPEED OF LEARNING]
Arabic	5336601	yes	to draw a connection between the sound and the written form [VISUAL ASSOCIATION]. it also helps if I miss how each sound is pronounced. [CORRECT PRODUCTION]
Arabic	5335700	yes	So I can differentiate between the sounds. [CORRECT COMPREHENSION]
Arabic	5331395	yes	لكي افهمها جيداً [CORRECT COMPREHENSION]
Arabic	5299874	yes	easier to remember [REMEMBERING]
Arabic	5298847	yes	Because if I don't, I sometimes mishear some sounds. The same is true in both languages and especially in songs [CORRECT COMPREHENSION]. Looking at orthography also helps with etymology and morphemes [VISUAL ASSOCIATION].
Arabic	5212730	yes	لمعرفة كيفية كتابتها [CORRECT PRODUCTION]
Arabic	5247103	yes	In English , there is a difference between the speling and the pronouncing of words, which make it importing to learn the the speling of a new word when hearing it .[DECODING SOUNDS AND LETTERS]
Arabic	5246620	yes	to make easier for me *[EXCL - INSUFFICIENT]*
Arabic	5155297	yes	to ramber the word [REMEMBERING]
Arabic	5196532	yes	mental connection between letters and how the sound [DECODING SOUNDS AND LETTERS]. make sure about the right sound. for example, bin or ban. to make sure i got the right word [CORRECT COMPREHENSION]
Arabic	5201553	yes	This way you remember it better. [REMEMBERING]
Arabic	5212731	no	ممكن أن أتشتت عن استماع الكلمة بالشكل الصحيح وأركز على شكها فقط [IT'S TOO MUCH TO PROCESS]

Arabic	5182923	yes	to know how to spell it [CORRECT PRODUCTION]
Arabic	5201938	yes	to remember the word [REMEMBERING] and know how to write it [CORRECT PRODUCTION]
Arabic	5182922	yes	لكي يسهل حفظها [MEMORISATION]
Arabic	5189512	no	not sure *[EXCL - INSUFFICIENT]*
Arabic	5155797	no	to make sure I pronounce it correctly [CORRECT PRODUCTION]
Arabic	5089630	yes	to remember them [REMEMBERING]
Arabic	5097676	yes	تثبت الكلمة اسرع [SPEED OF LEARNING]
Arabic	5029580	yes	reading the spelling of a word makes it easier to learn [CLARIFYING AND CONSOLIDATING NEW LANGUAGE]
Arabic	5082294	yes	to remember it [REMEMBERING]
Arabic	5018214	yes	for spelling purposes [CORRECT PRODUCTION] and also it makes it easier to remember the word [REMEMBERING]
Arabic	5003021	#N/A	#N/A
Arabic	5027737	yes	Seeing the spelling helps me remember the word better. [REMEMBERING]
Arabic	5010776	yes	To memorize it quickly [SPEED OF LEARNING]
Arabic	5047738	yes	ليسهل علي حفظها [MEMORISATION]
Arabic	5038591	no	لا اعرف *[EXCL - INSUFFICIENT]*
Arabic	5044345	yes	لتسهيل معرفتها [CLARIFYING AND CONSOLIDATING NEW LANGUAGE]
Arabic	5037277	yes	in order to memorize that word [MEMORISATION]
Arabic	5029573	yes	يسهل تذكرها [REMEMBERING]
Arabic	5010775	yes	to memorise the spelling and its pronunciation [MEMORISATION]
Arabic	5034149	yes	to help me connect the voice with the letters [DECODING SOUNDS AND LETTERS]
Arabic	5016049	yes	كي أكتبها [CORRECT PRODUCTION]

Arabic	5031281	yes	to memoriese the new word [MEMORISATION]
Arabic	5031282	yes	يسهل تذكرها [REMEMBERING]
Arabic	5029568	yes	ليسهل علي حفظها ومعرفة حروفها [MEMORISATION]
Arabic	5029577	yes	to easy learn *[EXCL - INSUFFICIENT]*
Arabic	5029572	yes	احفظها بشكل اسرع [SPEED OF LEARNING]
Arabic	5029579	yes	to make it easier to remember [REMEMBERING]
Arabic	5027663	yes	لأني بصرية [LEARNING STYLE]
Arabic	5027661	yes	to ermember the word [REMEMBERING]
Arabic	5027570	yes	because it will make it easier to understand how can i speak the word and how can i spell it [CORRECT PRODUCTION]. ~~a~~lso every time i see the picture i remember how it has been written this way i can choose fastly *[EXCL - ABOUT THE EXPERIMENT]*.
Arabic	5003012	yes	تساعدني على ربطها بالصورة و تخزينها في الذاكرة يكون اسهل و اسرع *[EXCL - ABOUT THE EXPERIMENT]*.
Arabic	5003019	no	i don`t need [IT'S TOO MUCH TO PROCESS]
Arabic	5010884	yes	لأن العين تحفظ ايضا [MEMORISATION] ولعدم القدرة على تمييز بعض الاحرف عند سماعها فصعبت التفريق بين نطق بعض الكلمات [CORRECT COMPREHENSION]
Arabic	5003023	yes	لكي اربط في ذهني الكلمة المكتوبة مع الصوت ... احفظ النطق و الكتابة معا في نفس الوقت [VISUAL ASSOCIATION]
Arabic	5010779	yes	لأكتبها و أقرئها صحيحة [CORRECT PRODUCTION]
Arabic	5003015	yes	To memorise it correctly [MEMORISATION]
Arabic	5000809	yes	لأنها تسهل عملية حفظها في الذاكرة المرئية [MEMORISATION] و ربطها مع الذاكرة السمعية [VISUALISATION]
Arabic	5000812	yes	to know more [CLARIFYING AND CONSOLIDATING NEW LANGAUGE]
Arabic	5000759	no	انا شخصية سمعية [PREVIOUS LANGUAGE LEARNING]
Arabic	5000407	yes	للتأكد من صحة ما استمعت إليه [CORRECT COMPREHENSION]
Arabic	4988182	yes	سهلة الحفظ [MEMORISATION]

Arabic	4986648	yes	make sure how it is prounanced [CORRECT PRODUCTION]
Arabic	4986649	yes	notice the difference between v and f better *[EXCL - ABOUT THE EXPERIMENT].*
Arabic	4986647	yes	حتى اتخيل شكل الكلمة [VISUALISATION]

Q7: In general, when learning new words do you think it is important to see the spelling the first time you hear it?
بشكل عام، عندما تتعلم كلمات جديدة هل من المهم عندك أن ترى املاء الكلمة عند الاستماع لها لأول مرة ؟
Q8: Why?
لماذا؟

Group	ID	Response [with tags]	
English	5304273	no	easier to hear and then understand how its written [SOUND IS MORE IMPORTANT/BENEFICIAL]
English	5305661	yes	easier way to distinguish or make a memory to a picture/object *[EXCL - ABOUT EXPERIMENT]*
English	5308502	yes	association [VISUAL ASSOCIATION]
English	5308657	yes	Because I'm quite a visual and textual learner and slightly deaf, so I rely more on what I see than what I hear. [LEARNING STYLE]
English	5308960	yes	Easily noticeable differences in spelling, allowed differentiation between words that did sound pretty similar *[EXLC - ABOUT EXPERIMENT]*
English	5312160	yes	Because otherwise I do not have the correct "image" of the word and may learn it distorted. In order to hear a word correctly I have to hear it repeated quite a few times. [CORRECT COMPREHENSION]
English	5312301	yes	To support learning *[EXCL - INSUFFICIENT]*
English	5312548	yes	Learning the spelling of a word allows you to use it in writing. I find it frustrating being unsure about the spelling of a word when I want to use it in writing. [CORRECT PRODUCTION]
English	5315514	yes	It helps reinforce it [CLARIFYING AND CONSOLIDATING NEW LANGUAGE]
English	5322227	yes	It helps visualise the word [VISUALISATION]

English	5323642	yes	Because I think it's so easy to mishear words. Especially in the flow of a conversation, and a new word jumps out to you. If I've been watching a TV program, and heard a new word, I'll often google it – and find I've spelt it incorrectly because I've misheard the pronunciation. Or the person speaking the word may have an accent, so things like vowels or inflections can get misinterpretted. The funniest instance of this is hearing new songs for the first time – I'm 98% guarenteed to sing several words/lines of the song wrong, because my brain fills in what it think's it's hearing with the first word it sounds like to me. But on googling the lyrics, I see what the artist is actually singing about [CORRECT COMPREHENSION]. So for me, I think seeing spelling/words written out, helps me to remember what I'm hearing [REMEMBERING], and helps me to correcty pronounce (or sing) the words and songs in the future. Though you do make people laugh when you say things wrong and that can be funny! [CORRECT PRODUCTION]
English	5327303	yes	Moslty so you know, or can at least have a good guess, on how to pronounce it [CORRECT PRODUCTION]
English	5328002	no	because hearing and seeing is enough [IT'S TOO MUCH TO PROCESS]
English	5330132	yes	helps you understand and remember ot corectly and hear the words are slightly different [CORRECT COMPREHENSION]
English	5332522	yes	it helps me sound it out loud [DECODING SOUNDS AND LETTERS], which helps me with pronunciation [CORRECT PRODUCTION] and word recognition [CORRECT COMPREHENSION], as well as with recall [REMEMBERING]
English	5332693	yes	Visual learner, my memory seems quite dependent on having something in front of me to help with learning [LEARNING STYLE]
English	5333575	yes	visual and sound connection [VISUAL ASSOCIATION]
English	5335169	yes	Because some letters sound similar, it re-enforces you are hearing the correct sound [CORRECT COMPREHENSION]
English	5336201	yes	because you can then visualise the word as well [VISUALISATION], and maybe notice patterns in spellings if it is similar to a word you already know [DECODING SOUNDS AND LETTERS]. You can also make sure you are pronouncing it correctly if you see it. (if you can read) [CORRECT PRODUCTION]

English	5337460	no	Maybe the first time it's important to try to focus only on the sounds, but by the second time, I do prefer to have the written version too.[SOUND IS MORE IMPORTANT/ BENEFICIAL]
English	5338550	yes	to confirm youve heard the sounds correctly [CORRECT COMPREHENSION]
English	5338600	yes	Allows you to see all the different letters in the word that you might miss through hearing [CORRECT COMPREHENSION]
English	5339584	yes	helps to remember [REMEMBERING]
English	5339906	yes	to learn it correctly from the beginning [CLARIFYING AND CONSOLIDATING NEW LANGUAGE]
English	5340413	yes	Actually, it depends, but the questionnaire doesn't allow for this answer... I learned some words better WITH spelling, and some words better WITHOUT the spelling, weird? *[EXCL - INSUFFICIENT]*
English	5341833	no	English spelling can be misleading (e.g. the second vowel sound in "Nasit" sounded more like the oo in "foot" to me). Phonetic spelling perhaps more helpful (although this also needs learning, of course!) [SPELLING DOESN'T REFLECT PRONUNCIATION]
English	5341870	yes	Because you can identify with the word easier when you see it written down in a text [CORRECT COMPREHENSION]
English	5343885	yes	To help visualise the word [VISUALISATION]
English	5344305	no	Because, that's how I learnt French. However, my spelling is awful so it probably is important to do so but it wasn't something that stopped me learning to speak. [PREVIOUS LANGUAGE LEARNING]
English	5353998	yes	to avoid mishearing [CORRECT COMPREHENSION]
English	5354235	yes	it hekps clarify the letters eg f or. v [CORRECT COMPREHENSION]
English	5356536	no	speaking and listening is the first skills in a new language [PREVIOUS LANGUAGE LEARNING]
English	5356537	yes	to ensure dont mishear, esoecially with similar letters, like v and f or m and n [CORRECT COMPREHENSION]
English	5357670	no	I think it interferes with my listening to see the word, whether in Arabic or English. [IT'S TOO MUCH TO PROCESS]

English	5359677	no	I focused more on sounds than spelling *[EXCL - ABOUT THE EXPERIMENT]*
English	5366758	yes	I find it difficult to recall the word, or understand it fully without seeing it spelt [CORRECT COMPREHENSION]. It provides a visual cue [VISUAL ASSOCIATION].
English	5366866	yes	No *[EXCL - INSUFFICIENT]*
English	5374409	yes	I personally visualise what I hear [VISUALISATION]
English	5374964	yes	if they sound similar to another word it's easier to distinguish [CORRECT COMPREHENSION]
English	5380819	no	no *[EXCL - INSUFFICIENT]*
English	5384030	yes	it links the image and sound to the picture *[EXCL - INSUFFICIENT]*
English	5384032	yes	i find it easier to read the word first before looking at the picture *[EXCL - INSUFFICIENT]*
English	5384657	no	learn by sound [SOUND IS MORE IMPORTANT/BENEFICIAL]
English	5409941	no	It depends on how the word the is learnt. If it is a speed learning when I'm trying to remember the sound of words, a picture associated with their meaning, and their spelling at once, it seems to overtask me. Word and picture, or sound and picture work, but all three together in a timed test is a bit much. [IT'S TOO MUCH TO PROCESS]
English	5417849	yes	Helps understand how a word is pronounced [CORRECT PRODUCTION]
English	5418540	yes	Seeing the word means you take the word in visually as well as through hearing it, making it easier to learn. [VISUAL ASSOCIATION]
English	5434242	no	If the focus is purely on sound, hearing and then repeating orally is more beneficial in consistently replicating and identifying the sound/word [SOUND IS MORE IMPORTANT/ BENEFICIAL]
English	5499381	no	not focused on written just the sound [SOUND IS MORE IMPORTANT/ BENEFICIAL]

English	5499420	yes	It helps you break down the different sounds/letters involved, especially if there are three or more syllables [DECODING SOUNDS AND LETTERS]. It prevents you learning it incorrectly incase you have mishead the pronounciation [CORRECT COMPREHENSION].
English	5499421	yes	I think I find it quite difficult to make the correct sounds when speaking, without first seeing the word written down (assuming it is written in characters I am familiar with) [CORRECT PRODUCTION]
English	5502013	yes	because i can attempt an answer phonetically [CORRECT PRODUCTION]
English	5503474	yes	It can give an indication of which sound you are lisening for in English - though not always [CORRECT COMPREHENSION]
English	5510526	yes	Actually maybe not the first time you hear them but I find it useful to see new words written down. I think it helps me to help me remember them. [REMEMBERING]
English	5521508	no	Spelling rarely reflects pronunciation in English and in the languages I have learnt. [SPELLING DOESN'T REFLECT PRONUNCIATION]
English	5557001	yes	It offers a visual cue from the very start that I can use to recall the word later. [VISUAL ASSOCIATION]
English	5586643	yes	soliifies understanding of hearing the word, like a confirmation [CORRECT COMPREHENSION]
English	5586843	yes	I think it is probably subconsciously helping - even though I dont think it helps me! *[EXCL - INSUFFICIENT]*
English	5587514	no	so many english words are spelt in different ways that dont always make it easy on first appearance [SPELLING DOESN'T REFLECT PRONUNCIATION]
English	5588274	yes	it helps me to understand what the word means [CORRECT COMPREHENSION]
English	5592006	yes	spellings provide a second sense to link the memory of the word to (sight as well as sound) [VISUAL ASSOCIATION] and can give you clues as to keaning which can also help memory [REMEMBERING].
English	5595978	no	Because I can't correlate the sound I'm hearing to the shape of the lettering in Arabic. It would be different if I had learnt the

			Arabic alphabet first maybe?? *[EXCL - ABOUT THE EXPERIMENT]*
English	5602966	yes	It helps to clarify which letters are used, and helps me to remember what the word is if I have seen it written as well as heard it [REMEMBERING]. Without seeing it written, it can be hard to make sense of the sounds being made if the sounds are new to me. [CORRECT COMPREHENSION]
English	5615580	yes	i guess it's so that i can match sounds to letters [DECODING SOUNDS AND LETTERS]
English	5620967	no	It can be a distraction [IT'S TOO MUCH TO PROCESS]
English	5621683	no	I can memorize a word and a meaning pretty quickly. The spelling adds another layer of memorization. [IT'S TOO MUCH TO PROCESS]
English	5627982	yes	Helps embed in uour brain [REMEMBERING]
English	5627983	no	its not always a good guide to pronunciation [SPELLING DOESN'T REFLECT PRONUNCIATION]
English	5629029	no	concentrating on the picture & the sound, seeing the spelling too is overload *[EXCL - ABOUT THE EXPERIMENT]*
English	5637584	yes	help make a connection [VISUAL ASSOCIATION]
English	5642980	yes	I find it easier to remember the word and place it with the object in my memory *[EXCL - ABOUT THE EXPERIMENT]*
English	5649032	yes	Seeing and hearing a word at the same time offers more than one way in to learn the word: to associate the sound of the word with how you conceptualise pronouncing the letters, and/or to associate the sound with the image of the letters (if they can't yet be read and pronounced). [VISUAL ASSOCIATION]
English	5655323	no	spelling in English can be confusing [SPELLING DOESN'T REFLECT PRONUNCIATION]
English	5659195	no	Because when you first learn a language, as a young child, its mostly done through listening [PREVIOUS LANGUAGE LEARNING]
English	5659326	no	In english the spelling doesn't always match up with pronunciation so it could make it difficult. [SPELLING DOESN'T REFLECT PRONUNCIATION]
English	5663887	yes	It makes it easier for me to learn, I am a pretty visual learner

			[LEARNING STYLE]
English	5664502	yes	I prefer to be able to see the spelling - or IPA if the sound-spelling correspondence is not 100% - in order to create a clearer mental representation in my mind [CLARIFYING AND CONSOLIDATING NEW LANGUAGE], and to better distinguish certain sounds from other sounds if there are similar sounds, or words with minimal pairs [CORRECT COMPREHENSION]
English	5666899	no	Words can be pronounced differently to how they look [SPELLING DOESN'T REFLECT PRONUNCIATION]
English	5667662	yes	So that you can begin learning how to recognise how certain letters are pronounced [DECODING SOUNDS AND LETTERS]
English	5670727	yes	it helps to register how i should pronounce the word correctly. If i always only listened to a person pronouncing it i may not say it right because the person speaking could have an accent and not clearly pronounce each sound that should be made when speaking clearly. Therefore without any spelling it would make me feel insecure about my speaking abilities and be worried my fluency would be affected by anothers mispronuncuation or natural talking habits. [CORRECT PRODUCTION]
English	5671003	yes	I am more of a visual learned than an aural learner and it is helpful to see the sound matched up with letters [LEARNING STYLE]
English	5672324	yes	Helps me to see the word and visualise it [VISUALISATION]
English	5674519	yes	helps visualise the word [VISUALISATION]
English	5676564	yes	to match the letters to the sounds [DECODING SOUNDS AND LETTERS]
English	5690497	yes	I guess maybe I learn visually better than hearing the words [LEARNING STYLE]. I find it difficult to distinguish between sounds. Pronunciation is always the most difficult part of learning a language for me [CORRECT PRODUCTION] and similarly hearing the difference between sounds that perhaps there is no difference in english is very difficult for me [CORRECT COMPREHENSION].
English	5691858	yes	To be able to say the words correctly [CORRECT PRODUCTION] and visualise the word [VISUALISATION]
English	5692078	yes	Visual reminder [VISUAL ASSOCIATION]

English	5692100	yes	It helps my memory[REMEMBERING]
English	5696781	yes	I think I sometimes linked the shape of a prominent letter with the object. For example the 'vadit' was like an upside down V. *[EXCL - ABOUT THE EXPERIMENT*
English	5704997	yes	It helps me to remember the word [REMEMBERING]
English	5706490	no	i feel the structure of the word is more important. number of syllables and the vowels that were used. [SOUND IS MORE IMPORTANT/BENEFICIAL]
English	5706491	no	When learning a new word, just hearing it rather than hearing it and seeing it requires more effort to retain therefore I think you'd be able to remember it more easily as more effort has gone into remembering it. [SOUND IS MORE IMPORTANT/ BENEFICIAL]
English	5719579	yes	That way you know for sure what the sounds are that make up the word, and that there isn't any pronunciation variation [CORRECT COMPREHENSION]
English	5732647	yes	This clarifies any misconception with sounds, especially when the words are similar [CORRECT COMPREHENSION]
English	5750023	no	Spelling doesn't always help with pronunciation, sometimes it's easier to listen and repeat without thinking about how to spell it [SPELLING DOENS'T REFLECT PRONUNCIATION]
English	5750994	yes	To build the concept of things in my brain [CLARIFYING AND CONSOLIDATING NEW LANGUAGE]
English	5751031	no	For me, I find oral fluency much easier to reach than written. I prefer to learn languages by speaking and hearing them, not writing (which I have done as well, with dead languages). [PREVIOUS LANGUAGE LEARNING]
English	5755297	yes	As mentioned, I am a more visual learner than auditory learner [LEARNING STYLE]. Being able to associate the words with the images helped a lot more *[EXCL - ABOUT THE EXPERIMENT]*
English	5755298	no	I don't think it's important that it's shown the *first* time [SOUND IS MORE IMPORTANT/BENEFICIAL], but it should be shown eventually and the different sounds should be broken down step-by-step (at least for beginners of the language)
English	5755299	yes	hints about etymology, pluralization/declendion and pronunciation [VISUAL ASSOCIATION]

English	5755300	yes	Because it helps for a learner to recognise patterns within the spelling system and how they interact with the sounds [DECODING SOUNDS AND LETTERS]
English	5755301	yes	in a natural learning environment, there could be sound distractions which make it harder to hear the correcr pronunciation of the word. Seeing the spelling would help the listener clearly hear the pronunciation [CORRECT COMPREHENSION
English	5756084	yes	Spelling gives an anchor to schematize the sounds you hear [CORRECT COMPREHENSION]
English	5756085	yes	I am a text-based learner, I believe [LEARNING STYLE]
English	5758767	no	I'm not sure it makes a difference the first time because the matching exercise is quite easy, but it's helpful if you're trying to cement it to learn it for the future *[EXCL - ABOUT THE EXPERIMENT]*
English	5759398	yes	to correct hearing errors [CORRECT COMPREHENSION]
English	5759735	no	hearing is most helpful first [SOUND IS MORE IMPORTANT/ BENEFICIAL]. written and hering the word is more to process [IT'S TOO MUCH TO PROCESS]
English	5762777	yes	better to know which consonant is being spoken right away to remember it correctly [CORRECT COMPREHENSION]
English	5765135	yes	because it can be important to get an idea of what the letters sound like which for me at least makes it a lot easier to remember them especially if they only have one letter difference [DECODING SOUNDS AND LETTERS]
English	5765136	yes	It makes learning to read in the language much easier. [DECODING SOUNDS AND LETTERS]
English	5780967	yes	helps me map the word better in my vocab [VISUAL ASSOCIATION] im a speech therapist if you want to know more about what i mean by that
English	5831108	yes	helps you be able to pronounce it and realise what word you are saying [CORRECT PRODUCTION]
English	5884960	no	Na *[EXCL - INSUFFICIENT]*
English	5909329	yes	additional aid to memory [REMEMBERING]

English	5909763	yes	It is helpful for people who are more visual learners [LEARNING STYLE]. Also helpful when the language being learnt has a different alphabet, so trying to learn phoentically is important [DECODING SOUNDS AND LETTERS]
English	5914367	no	Easier to concentrate on one aspect of the word first. ie just hearing it or just seeing it [IT'S TOO MUCH TO PROCESS]
English	5998960	yes	If you have a visual indicator of the word (via its spelling), you can remember it more easily [REMEMBERING]. You can make stronger connections between the word, its spelling and the concept it represents [VISUAL ASSOCIATION].
English	6029761	yes	If you are a very visual learner [LEARNING STYLE], you can picture the word spelling when you say it [VISUALISATION]

Appendix XVI: Coder guidelines for Q13

The different categories are outlined in the tables below, accompanied by definitions and examples. There are 6 main categories and 7 subcategories to consider. At the end of the document, there is a random sample of 20% of participant responses which should be coded according to the coding guidelines provided.

The recommended steps are:

1. Discard responses that are too vague to be coded or are not directly related to the category definition or research question.
2. Go back through material and follow the coding guide to apply the more specific categories.
3. Make notes of any categories that you feel do not fit well or responses that may require the formulation of a new or different category.

Of course, if you have any questions, don't hesitate to ask!

Coding Arabic responses

Every unit of analysis was assessed for its relevance to the category definition and coded with the level of abstraction in mind. This meant that units of analysis must include reference to an action chosen and purposefully applied to the task of learning the (phonological) form of new words, with particular attention paid to the influence of modality of input. To focus on the active nature of strategy application, all categories were formed with the present participle of a verb (Oxford, 2016). These categories can apply broadly or in specific reference to English or Arabic script. L1 Arabic-speaking participants could choose whether to respond in English or Arabic and the original, untranslated and unchanged responses have been provided for coding. All categories are English, but keywords and examples have been provided in Arabic.

Table XVI.1: L1 Arabic coding guidelines for language learning strategies used during the task

Category	Definition	Examples
Remembering (broadly)	Broad focus on trying to remember the word, without specifying further. **Arabic keywords** تذكر, حفظ	"At this point my remembering" "بالاعتماد على الذاكرة فقط"
Focusing… • On sounds	Focus on remembering or paying attention to the sounds of a word, in terms of perception, recognition or production. **Arabic keywords** تذكر/ حفظ/ التركيز ما سمعته/الصوت/ النطق/ اللفظ	"I tried to remember the voice saying it" "أحاول تذكر استماعي للكلمة"
• On letters	Focus on remembering or paying attention to the letters or spelling of a word, usually the first letter. **Arabic keywords** تذكر/ حفظ/ التركيز الحرف الأول, الإملائية	"From remembering the first letter" "بالحرف الأول من كل كلمة"
• On images	Focus on remembering or paying attention to the images, as a whole of a detail of the picture. **Arabic keywords** تذكر/ حفظ/ التركيز, الصورة, الشكل	"concentrating on the visual image and the colors in the image" "من خلال حفظ الصورة"
Associating (broadly)	Broad focus on associating different elements of the input, usually sound and image. It may include spelling. It may be connected to memory/ remembering. It does not specify further. **Arabic keywords** اربط	"link between what i hear and see" "اربط الصورة بالصوت و الكلمة"

• With something familiar	Specific focus on associating new words with familiar words in a known language or object. Can involve creating stories, contextual sentences or movement to facilitate the association. **Arabic keywords** ربطتها بكلمات اعرفها, مشابهة في النطق, بشيء اعرفه	"By sound associations with word I already know." "بربطها باشياء اخرى في الحياة اليومية مثلا كلمة fumel تبدا بحرف الفاء و تشبه الوردة " flower " "
• With visual detail	Specific focus on associating the word with a specific detail from the image, e.g. colour, shape, texture etc. **Arabic keywords** بالألوان	"concentrating on the visual image and the colors in the image."
• With letter shapes	Specific focus on associating the image with the shape of the first letter or shape of the written word. **Arabic keywords** ربط الحروف بالشكل	"Linking shapes for the letters with the object – eg Mackem had 3 legs, like an M" "الاعتماد على الذاكرة فقط وربط الحروف بالشكل"
Repeating	Broadly states the use of repetition out loud or in the mind/silently. It may reference memory but does not specify further **Arabic keywords** أحاول نطق, التكرار	"I repeated them to myself out loud after hearing and or seeing them." "Repeated the word in my head" "أحاول نطق الكلمة" "بعض الكلمات ترسخت في ذاكرتي بسبب التكرار"
Comparing/ contrasting	Looking for similarities and differences between sounds, letters, images and words in general. May include grouping or finding patterns. May refer to memory. **Arabic keywords** الكلمات التي تختلف, قارن	" I attempted to memorise the first letter which was often the difference between the two similar sounds eg the voiced vammell and voiceless fammell"

		"في مرحلة التعليم، كنت أحاول ربط الكلمات التي تختلف في صوت واحد فقط، مثلا كلمة "ناست" و"ماست" تختلفان في الصوت الأول فقط، وهذا سهل علي التعلم"
Visualising	Creating mental images/pictures based on the presented input. May be used to combine image, spelling and/or audio in memory. **Arabic keywords** اتخيل /إنشاء صورة, العقلية/ في ذهن	"visualizing how the words sound incoproated with how they're spelt" "Mental image"
Evaluating learning	Evaluating success of learning in general or particular strategies. May include switching strategies based on evaluation. **Arabic Keywords** ساعد أكثر, ناجحة, أفضل/أسوأ	"I also sometimes used the shape of the item e.g. macum looked a bit like an "m" so I used that to remember. But definitely the repetition helped more than anything." "بمحاولة تذكر نطقها ولكن شعرت بتداخل بين كلمات علقت بذهني غير مرتبطة بالصور الصحيحه"
Directing learning	Structuring the approach to the task through directing attention, going beyond the task, or choosing how to engage with input. Often mentioned in relation to cognitive strategies. **Arabic keywords** التركيز المتعمد, إيقاع, منظم, لعبة	"I said them out loud twice after each was shown, I then repeated the names once more before the testing phase. I almost had it in a rhythm that kept the learning time very fast-paced but consistent." "لم أركز كثيرا على الحروف المتشابهة كي لا أتشتت كنت أتعامل مع كل كلمة على حدى"

Table XVI.2: L1 Arabic example of coding approach

ID	**How did you try to learn or remember the words?** **كيف حاولت تعلم أو تذكر الكلمات ؟**	**codes**
5370516	في مرحلة التعليم، كنت أحاول ربط الكلمات التي تختلف في صوت واحد فقط، مثلا كلمة "ناست" و"ماست" تختلفان في الصوت الأول فقط، وهذا سهل علي التعلم	comparing/ contrasting
5247103	link it with somthing i know	Associating with something familiar
5357676	احيانا من شكل الصورة احاول اربط شكل الصورة بشي عربي مثل الشي الازرق .. فقد ربطه بكلمة (نجم) بالعربية لان له أذرع	Associating with something familiar
5010884	بمحاولة تذكر نطقها ولكن شعرت بتداخل بين كلمات علقت بذهني غير مرتبطة بالصور الصحيحه	Focusing on sounds - Evaluating learning
5189512	close my eyes and try to remember	Directing learning - Remembering

Table XVI.3: L1 Arabic 20% random sample for 2nd coder

ID	**How did you try to learn or remember the words?** **كيف حاولت تعلم أو تذكر الكلمات ؟**	**codes**
5155797	connecting it to the picture	
5029587	recall	
5155297	first liter	
5339069	Depending on my memory of what I heard,	
5196532	I made visual picture. For example, I remeber one of the words buy connecting it to camel:) 🐪 because the rythem of the create word is similar. The last picture has legs that form letter m. So, i remeber that it starts with /m/	
5029568	من الذاكرة	
5034149	from remembering the picture	
5027570	i tried to think about how that person say it and guessed the spelling upon my experience	
5003019	make relation with the shape	
5003023	" تبدا بحرف الفاء و تشبه الوردة fumel بربطها باشياء اخرى في الحياة اليومية مثلا كلمة flower"	
5339892	بالاعتماد على الذاكرة فقط وربط الحروف بالشكل	

5359087	أربط بين الصوت والشكل، مثلا هناك شكل ناسيت وهو عبارة عن إناء مثقوب، فربطت ذلك بفعل نسيت في العربيةوالرابط مجازي وهو ثقوب الذاكرة	
5359752	ربطت صوت الكلمة بصورة في ذهني	
5337115	اتذكر نطقها ومشاهدتها لكن ليست كلها مكتوبة	
5331395	اتخيل اني حفظتها	
5229722	I tried to create a linkage between how the word sounds and how the image looks.	
5337166	تذكر النطق	
5031281	through using linking	
5553283	عن طريق التركيز فى الصوت والصورة	
5082295	من تذكر كيفيه استماعها	
5027737	I tried to connect the shape of the thing to something similar in real life.	
5029579	by remembering the sounds of these words	
5011969	ارتباطها بالصور	
5341799	I linkd the shape of the object with a word or letter. for example the last shape in the previous question looks like M. The Madas shape remined me with "madas" which means shoes in Makkah & Jeddah people slang. there alos an object looks like fallopian tube.	

Coding English responses

Every unit of analysis was assessed for its relevance to the category definition and coded with the level of abstraction in mind. This meant that units of analysis must include reference to an action chosen and purposefully applied to the task of learning the (phonological) form of new words, with particular attention paid to the influence of modality of input. To focus on the active nature of strategy application, all categories were formed with the present participle of a verb (Oxford, 2016). These categories can apply broadly or in specific reference to English or Arabic script.

Table XVI.4: L1 English coding guidelines for language learning strategies used during the task

Category	**Definition**	**Examples**
Remembering (broadly)	Broad focus on trying to remember the word, without specifying further. **English keywords** Remember, memorise, recall	“At this point my remembering”
Focusing… ● On sounds ● On letters ● On images	 Focus on remembering or paying attention to the sounds of a word, in terms of perception, recognition or production. **English keywords** remember/ memorise/ focus, sound/ voice/ pronunciation Focus on remembering or paying attention to the letters or spelling of a word, usually the first letter. **English keywords** remember/ focus; first/last letter, spelling Focus on remembering or paying attention to the images, as a whole of a detail of the picture. **English keywords** concentrate/ focus/ study / remember, image/ picture/ object/ shape	 “I tried to remember the voice saying it” "From remembering the first letter" "concentrating on the visual image and the colors in the image"
Associating (broadly)	Broad focus on associating different elements of the input, usually sound and image. It may include spelling. It may be connected to memory/ remembering. It does not specify further. **English keywords** Linking, connecting, relating, sounds like, looks like	"link between what i hear and see”

• With something familiar • With visual detail • With letter shapes	Specific focus on associating new words with familiar words in a known language or object. Can involve creating stories, contextual sentences or movement to facilitate the association. **English keywords** Looks like, sounds like, reminds me of, associate with, familiar Specific focus on associating the word with a specific detail from the image, e.g. colour, shape, texture etc. **English keywords** Colours, distinctive features, aspect of shape Specific focus on associating the image with the shape of the first letter or shape of the written word. **English keywords** Looked like letter, make a mental shape	"By sound associations with word I already know." "concentrating on the visual image and the colors in the image." "Linking shapes for the letters with the object – eg Mackem had 3 legs, like an M"
Repeating	Broadly states the use of repetition out loud or in the mind/silently. It may reference memory but does not specify further **English keywords** Repeat	"I repeated them to myself out loud after hearing and or seeing them." "Repeated the word in my head"
Comparing/ contrasting	Looking for similarities and differences between sounds, letters, images and words in general. May include grouping or finding patterns. May refer to memory. **English keywords** Comparing, contrasting, differences	" I attempted to memorise the first letter which was often the difference between the two similar sounds eg the voiced vammell and voiceless fammell"
Visualising	Creating mental images/pictures based on the presented input. May be used to combine image, spelling and/or audio in memory.	"visualizing how the words sound incoproated with how they're spelt"

	English keywords Visualise, create/imagine picture/image, mind/mental.	“Mental image”
Evaluating learning	Evaluating success of learning in general or particular strategies. May include switching strategies based on evaluation. **English keywords** Helped more, successful, better/worse	“I also sometimes used the shape of the item e.g. macum looked a bit like an "m" so I used that to remember. But definitely the repetition helped more than anything.”
Directing learning	Structuring the approach to the task through directing attention, going beyond the task, or choosing how to engage with input. Often mentioned in relation to cognitive strategies. **English keywords** Intentionally, rhythm, structure, game, focus	“I said them out loud twice after each was shown, I then repeated the names once more before the testing phase. I almost had it in a rhythm that kept the learning time very fast-paced but consistent.”

Table XVI.5: L1 English example of coding approach

ID	How did you try to learn or remember the words?	codes
5333573	Was easier to remember the second one, and when asked to identify the picture, would hear either the second one, or if not, it had to be the first. The 2 pics which were not part of the test were discarded in my head, before making the judgement.	Directing learning
5602966	I sometimes tried to associate a letter with what I was seeing, e.g. was there a 'v' shape in any of the pictures that had a 'v' in the word? I also tried to assign some meaning to the object I was seeing which were all random objects to me. E.g. with the word famil (if I've rememebr that right), it looked a bit like an object that could create a funnel if inserted somewhere!	Associating with letter shapes Associating with something familiar
5692100	Said them out loud	Repeating
5620967	Spoke them out loud, repeated the words in my head. Focused on the first letter and sound.	Repeating Focusing on letters
5674519	associated with an english word that the picture made me think of	Associating with something familiar

Table XVI.6: L1 English 20% random sample for 2nd coder

ID	How did you try to learn or remember the words?	codes
5340413	As explained above, I sometimes associated them with something that looks or sounds familiar (English word).	
5751031	focusing on the consonants, especially those at the beginning of each syllable. I also realized you were mostly using words with minimal pairs (f/v, m/n) so focused on those especially.	
5756084	I focused on the contrasts in sounds	
5305661	Take the first letter and apply to another word that was relevant to the image	
5649032	By making connections between the sound of the word (more than its spelling - but thinking about what letters I thought I was hearing in it) and the image of the object the word described.	
5339584	sounds and colours compared to the other example	
5704997	I picked out a part of the image that i felt related to the word (the words with a V at the start i looked for a V shape in the object)	

5696781	Linking shapes for the letters with the object – eg Mackem had 3 legs, like an M	
5732647	Tried to link the new word image to a more familiar object or concept	
5354235	i pictured the initial letter in the picture	
5374964	by associating the pictures with more familiar objects	
5755301	I sounded out the words as I heard them and saw the pictures	
5434242	I said them out loud twice after each was shown, I then repeated the names once more before the testing phase. I almost had it in a rhythm that kept the learning time very fast-paced but consistent. I did, however, think that the learning phase was the timed game and so was intentionally learning them quickly rather than learning them well.	
5356537	looking at the shape to prompt / making up clues	
5831108	try to associate the made up words with what i thpught it kind of looked like	
5308502	made up stories about them	
5755297	By coming up with acronyms with the first letter in the word to associate it with the image.	
5499381	just from the sound and used the game like a memory game with pictures	
5755298	association of the sounds with the image displayed by repeating the word and pointing at the image	
5304273	remembering the endings	
5672324	Pneumonic device	
5332693	Repeating in my head what was said	
5690497	I repeated them in my head when the picture was shown. I also tried to link the words to the images in my head.	
5756085	Graft sound onto object	

Appendix XVII: Intercoder reliability calculations Q13

This document outlines the procedure to calculate Krippendorff's alpha, to assess intercoder agreement between three coders for both L1 English and L1 Arabic data in R (irr package). Relevant code and output are included below, adapted from Rmd files. First, it is necessary to ensure that the data is in the correct long form before loading into R (see table XVII.1). This is then reshaped in R to create new data frames for each variable with the content by coder. This was performed with the data for Q13, with the data from both language groups.

Table XVII.1 Example of data format before running R script

Participant_ID	Response	Content_ID	Coder_ID	Specific_code
1	xxxxx	1	A	Remembering
2	xxxxx	2	A	Focusing on letters
1	xxxxx	1	B	Remembering
2	xxxxx	2	B	Focusing on letters
1	xxxxx	1	C	Remembering
2	xxxxx	2	C	Focusing on letters

```
#Arabic coders language learning strategies

ar_specific_3 <- select(irr_3_ar, content_ID, coder_ID, specific_code)
ar_specific_3 <- pivot_wider(ar_specific_3, id_cols = coder_ID,
names_from = content_ID, values_from = specific_code)
ar_specific_3 <- select(ar_specific_3, -coder_ID)

#English coders language learning strategies

en_specific_3 <- select(irr_3_en, content_ID, coder_ID, specific_code)
en_specific_3 <- pivot_wider(en_specific_3, id_cols = coder_ID,
names_from = content_ID, values_from = specific_code)
en_specific_3 <- select(en_specific_3, -coder_ID)
```

In order to calculate Krippendorff's alpha for each variable, the data frames are transformed into matrices and reliability is calculated using the "kripp.alpha" function of the "irr" package. The output for the intercoder reliability for the categories within each variable is reported in table XVII.2.

```
ar_specific_3 <- as.matrix(ar_specific_3)
kripp.alpha(ar_specific_3, method = "nominal")

en_specific_3 <- as.matrix(en_specific_3)
kripp.alpha(en_specific_3, method = "nominal")
```

Table XVII.2: Krippendorff's alpha (α) assessing intercoder reliability of strategy categories, by L1

Category agreement (α)	**L1 English (n=3)**	**L1 Arabic (n=3)**
Q13 specific	0.84 (n=30)	0.71 (n=29)

Appendix XVIII: Coded questionnaire responses for Q13

Q 13: How did you try to learn or remember the words? كيف حاولت تعلم أو تذكر الكلمات ؟		
Group	**ID**	**Response [with tags]**
Arabic	4986647	ربط شكل اللعبة مع الكلمة. مثلا كلمة فامس شكل اللعبة شخص رافع يديه للأعلى. فامس من فيموس تعني مشهور وهذا المشهور سعيد و رافع يديه [ASSOCIATING WITH SOMETHING FAMILIAR]
Arabic	4986648	nadis looks like nest [ASSOCIATING WITH SOMETHING FAMILIAR]
Arabic	4986649	Guess *[EXCL- NO STRATEGY]*
Arabic	4988182	بالشكل [FOCUSING ON IMAGES] و الصوت [FOCUSING ON SOUNDS]
Arabic	5000407	النطق [FOCUSING ON SOUNDS]
Arabic	5000759	بالألوان [ASSOCIATING WITH VISUAL DETAIL]
Arabic	5000809	بشكل تلقائي [*EXCL- NO STRATEGY*]
Arabic	5000812	just I remembered what I have been listen [FOCUSING ON SOUNDS]
Arabic	5000814	by listening and repeating [REPEATING]
Arabic	5003012	ربط الصورة بالصوت [ASSOCIATING]
Arabic	5003015	from the pronouciation [FOCUSING ON SOUNDS]
Arabic	5003019	make relation with the shape [ASSOCIATING]
Arabic	5003021	N/A#
Arabic	5003023	بربطها باشياء اخرى في الحياة اليومية مثلا كلمة fumel تبدا بحرف الفاء و تشبه الوردة " flower" [[ASSOCIATING WITH SOMETHING FAMILIAR
Arabic	5010775	through hrearing them [FOCUSING ON SOUNDS]
Arabic	5010776	I couldnt remember thtem *[EXCL- NO STRATEGY]*
Arabic	5010779	لا أعلم *[EXCL- NO STRATEGY]*
Arabic	5010884	بمحاولة تذكر نطقها [FOCUSING ON SOUNDS] ولكن شعرت بتداخل بين كلمات علقت بذهني غير مرتبطة بالصور الصحيحه [EVALUATING LEARNING]

Arabic	5011969	ارتباطها بالصور[ASSOCIATING]
Arabic	5015277	i was trying to Link the word with something familiar to me using the picture [ASSOCIATING WITH SOMETHING FAMILIAR]
Arabic	5016049	حفظت الصوت [FOCUSING ON SOUNDS]
Arabic	5018214	i just guessed *[EXCL - NO STRATEGY]*
Arabic	5027570	i tried to think about how that person say it [FOCUSING ON SOUNDS]and guessed the spelling upon my experience [FOCUSING ON LETTERS]
Arabic	5027661	try to remember the sounds [FOCUSING ON SOUNDS]
Arabic	5027663	بالصوت [FOCUSING ON SOUNDS]
Arabic	5027737	I tried to connect the shape of the thing to something similar in real life. [ASSOCIATING WITH SOMETHING FAMILIAR]
Arabic	5029568	من الذاكرة [REMEMBERING]
Arabic	5029572	ربطتها بكلمات اعرفها [ASSOCIATING WITH SOMETHING FAMILIAR]
Arabic	5029573	لا استطيع تذكرها *[EXCL - NO STRATEGY]*
Arabic	5029577	Say *[EXCL - INSUFFICIENT]*
Arabic	5029579	by remembering the sounds of these words [FOCUSING ON SOUNDS]
Arabic	5029580	reading them in english and likning the sound to something I know (ex: makm = my laptop is Mac IOS) [ASSOCIATING WITH SOMETHING FAMILIAR]
Arabic	5029581	link between what i hear and see [ASSOCIATING]
Arabic	5029587	Recall [REMEMBERING]
Arabic	5031281	through using linking [ASSOCIATING]
Arabic	5031282	عن طريق ربطها بكلمات مشابهة في النطق لكامات عربية [ASSOCIATING WITH SOMETHING FAMILIAR]
Arabic	5034149	from remembering the picture [FOCUSING ON IMAGES]
Arabic	5037046	اربط الصورة بالصوت و الكلمة [ASSOCIATING]

Arabic	5037277	I was trying to remember the spelling.. I mean the sound 🔊 [FOCUSING ON THE SOUNDS]
Arabic	5038591	ربطتها بكلمات عربية [ASSOCIATING WITH SOMETHING FAMILIAR]
Arabic	5044345	تذكر ما استمعت له [FOCUSING ON SOUNDS]
Arabic	5047738	بالتذكر [REMEMBERING]
Arabic	5059373	i try to link the words or letters to already known objects [ASSOCIATING WITH SOMETHING FAMILIAR] or colors [ASSOCIATING WITH VISUAL DETAIL]
Arabic	5082294	by associating the object with a similar word [ASSOCIATING WITH SOMETHING FAMILIAR]
Arabic	5082295	من تذكر كيفيه استماعها [FOCUSING ON SOUNDS]
Arabic	5089629	N/A#
Arabic	5089630	I tried to remember the sound of them [FOCUSING ON SOUNDS]
Arabic	5097676	بالاستماع [FOCUSING ON SOUNDS]
Arabic	5155297	first liter [FOCUSING ON LETTERS]
Arabic	5155797	connecting it to the picture [ASSOCIATING]
Arabic	5182922	حاولت تذكر الصوت الذي سمعته [FOCUSING ON SOUNDS]
Arabic	5182923	the photos helped me to remember the first one [FOCUSING ON IMAGES]
Arabic	5189512	close my eyes and try to remember [DIRECTING LEARNING]
Arabic	5196532	I made visual picture [VISUALISING]. For example, I remeber one of the words buy connecting it to camel:) 🐪 because the rythem of the create word is similar [ASSOCIATING WITH SOMETHING FAMILIAR]. The last picture has legs that form letter m. So, i remeber that it starts with /m/ [ASSOCIATING WITH LETTER SHAPES]
Arabic	5201553	by continuous listening [FOCUSING ON SOUND]
Arabic	5201938	Jus what I remember from the gam [REMEMBERING]

Arabic	5212730	بالذاكرة. تذكر اول حرف [FOCUSING ON LETTERS]
Arabic	5212731	تذكرت صوتها [FOCUSING ON SOUNDS]
Arabic	5229722	I tried to create a linkage between how the word sounds and how the image looks. [ASSOCIATING]
Arabic	5246620	is very difficult *[EXCL - NO STRATEGY]*
Arabic	5247103	link it with somthing i know [ASSOCIATING WITH SOMETHING FAMILIAR]
Arabic	5298847	I split words in two syllables and try to remember the onset of each [FOCUSING ON SOUNDS]
Arabic	5299874	thier pics [FOCUSING ON IMAGES]
Arabic	5331395	اتخيل اني حفظتها [VISUALISING]
Arabic	5335700	Mental image. [VISUALISING]
Arabic	5335791	بالشكل الاملائي [FOCUSING ON LETTERS] او بربط الكلمة بشيء اعرفه مسبقا [ASSOCIATING WITH SOMETHING FAMILIAR]
Arabic	5336078	by linking the shape of the object with the sounds [ASSOCIATING]
Arabic	5336140	sometimes repeating the word [REPEATING] but not always(this failed me) [EVALUATING] + creating an image for the pic+word... ex the word famis for me seemed like a famous person who held his\her hand high.. trying possibly to wave to the fans? [ASSOCIATING WITH SOMETHING FAMILIAR]]other than that, some images were easier to remember because of the color(i.e., the orange thingy). [ASSOCIATING WITH VISUAL DETAIL]
Arabic	5336601	i got confused if I should write N or M in both makim and nakim and likewise with vadit and fadit. /v/ & /f/ *[EXCL - NO STRATEGY]*
Arabic	5336602	N/A#
Arabic	5336722	Connecting the shape to something in real world. e.g. those that are called 'famo' or 'vamo', I try to connect the word heard which is a bit similar to "فم" in Arabic saying the shape of the those pictures can be entered inside the 'mouth=فم". Another example, the shape that's called 'famis' ...I say it looks like a unique flower..so I remind myself that it's

		famis "like or driven from the English word famous". So in short, try to connect the new words to known words in either language Arabic or English. [ASSOCIATING WITH SOMETHING FAMILIAR]
Arabic	5337014	Guessing 😁 *[EXCL - NO STRATEGY]*
Arabic	5337115	اتذكر نطقها [FOCUSING ON SOUNDS] ومشاهدتها لكن ليست كلها مكتوبة [FOCUSING ON IMAGES]
Arabic	5337166	تذكر النطق [FOCUSING ON SOUNDS]
Arabic	5337282	i could not remember most of it [EVALUATING LEARNING] it was a really difficult exercise even i tried to repeat each words by my self [REPEATING]
Arabic	5337506	بالصوت [FOCUSING ON SOUNDS] والصوره [FOCUSING ON IMAGES]
Arabic	5337579	رؤية الصورة والكتابة والاستماع *[EXCL - MAKING NOTES]*
Arabic	5338152	بالتخمين *[EXCL - NO STRATEGY]*
Arabic	5338565	remember the sound [FOCUSING ON SOUNDS]
Arabic	5338811	#N/A
Arabic	5338948	خلال التخمين *[EXCL - NO STRATEGY]*
Arabic	5339069	Depending on my memory of what I heard, [FOCUSING ON SOUNDS]
Arabic	5339892	بالاعتماد على الذاكرة فقط [REMEMBERING] وربط الحروف بالشكل [ASSOCIATING WITH LETTER SHAPES]
Arabic	5341797	بعض الكلمات ترسخت في ذاكرتي بسبب التكرار [REPEATING]
Arabic	5341799	I linkd the shape of the object with a word or letter. for example the last shape in the previous question looks like M. [ASSOCIATING WITH LETTER SHAPES] The Madas shape remined me with "madas" which means shoes in Makkah & Jeddah people slang. there alos an object looks like fallopian tube. [ASSOCIATING WITH SOMETHING FAMILIAR]
Arabic	5341802	بالحرف الأول من كل كلمة [FOCUSING ON LETTERS]
Arabic	5342157	تخمين *[EXCL - NO STRATEGY]*

Arabic	5357676	احيانا من شكل الصورة احاول اربط شكل الصورة بشي عربي مثل الشي الازرق فقد ربطه بكملة (نجم) بالعربية لان له أذرع .. [ASSOCIATING WITH SOMETHING FAMILIAR]
Arabic	5359087	أربط بين الصوت والشكل، مثلا هناك شكل ناسيت وهو عبارة عن إناء مثقوب، فربطت ذلك بفعل نسيت في العربيةوالرابط مجازي وهو ثقوب الذاكرة [ASSOCIATING WITH SOMETHING FAMILIAR]
Arabic	5359091	تذكر الصوت [FOCUSING ON SOUNDS]
Arabic	5359092	لم احاول [*EXCL - NO STRATEGY*]
Arabic	5359752	ربطت صوت الكلمة بصورة في ذهني [ASSOCIATING]
Arabic	5365353	Memory [REMEMBERING]
Arabic	5369858	تخمين لانني لم اركز اطاقا على كتابتها في البداية عندما تم عرضها [*EXCL - MAKING NOTES*]
Arabic	5370466	أحاول تذكر استماعي للكلمة [FOCUSING ON SOUNDS]
Arabic	5370467	من خلال تذكر ما سمعته [FOCUSING ON SOUNDS]
Arabic	5370516	في مرحلة التعليم، كنت أحاول ربط الكلمات التي تختلف في صوت واحد فقط، مثلا كلمة "ناست" و"ماست" تختلفان في الصوت الأول فقط، وهذا سهل علي التعلم [COMPARING/CONTRASTING]
Arabic	5378097	الكتابة أثناء الاستماع [*EXCL - MAKING NOTES*]
Arabic	5384016	من الصوت [FOCUSING ON SOUNDS]
Arabic	5439128	مجرد تذكر بين ماسمعت والصورة [ASSOCIATING]
Arabic	5442543	ربط الكلمة بالشكل حتى لو بشكل بعيد جداا مثلا famis شكلها يشبه الجهاز التناسلي للأنثى —> female famist nackle تشبه مخدة الرقبة neck [ASSOCIATING WITH SOMETHING FAMILIAR]
Arabic	5450539	ربط الاشكل وتشابه الكلمات [ASSOCIATING]
Arabic	5452186	from remembering the first letter [FOCUSING ON LETTERS]
Arabic	5462535	at this point my remembering [REMEMBERING]
Arabic	5499647	بمحاولة تذكر لفظ تلك الكلمات [FOCUSING ON SOUNDS]

Arabic	5499711	Cockly *[EXCL - INSUFFICIENT]*
Arabic	5499774	I didn't remember *[EXCL - NO STRATEGY]*
Arabic	5500029	not seriously *[EXCL - NO STRATEGY]*
Arabic	5500050	with that words. not seriously *[EXCL - NO STRATEGY]*
Arabic	5501560	I think you gave me the words that I did not select the correct answers in the beginning so I have no link to remember *[EXCL - NO STRATEGY]*
Arabic	5553283	عن طريق التركيز فى الصوت [FOCUSING ON SOUNDS] والصورة [FOCUSING ON IMAGES]
Arabic	5607775	Remember to write *[EXCL - MAKING NOTES]*
Arabic	5622524	من خلال حفظ الصورة [FOCUSING ON IMAGES]
Arabic	5730442	بالربط المنطقي بين الشكل والكلمة [ASSOCIATING]، ولم أركز كثيرا على الحروف المتشابهة كي لا أتشتت كنت أتعامل مع كل كلمة على حدى [DIRECTING LEARNING] ركزت على الصوت [FOCUSING ON SOUNDS]، في بعض الأحيان كنت أتعمد تغطية الكتابة بأصابعي وأحاول نطق الكلمة قبل الصوت المسجل [REPEATING]

Q 13: How did you try to learn or remember the words? **كيف حاولت تعلم أو تذكر الكلمات ؟**		
Group	**ID**	**Response [with tags]**
English	5304273	remembering the endings [FOCUSING ON SOUNDS]
English	5305661	Take the first letter and apply to another word that was relevant to the image [ASSOCIATING WITH SOMETHING FAMILIAR]
English	5308502	made up stories about them [ASSOCIATING]
English	5308657	association of sound and picture. [ASSOCIATING]
English	5308960	not too sure, was only really thinking about the first exercise, which was a little more short-term memory and became a 50/50 right/wrong [*EXCL - INSUFFICIENT*]
English	5312160	By sound associations with word I already know.[ASSOCIATING WITH SOMETHING FAMILIAR]

English	5312301	Using the images [FOCUSING ON IMAGES]
English	5312548	Usually with some kind of mental association (e.g. 'famis' sounds like 'famous' and the picture looked like a person raising their arms in triumph) [ASSOCIATING WITH SOMETHING FAMILIAR] and repetition out loud [REPEATING]
English	5315514	by the first letter [FOCUSING ON LETTERS]. Repeating it [REPEATING]
English	5322227	Matching the audio sound with picture [ASSOCIATING]
English	5323642	By repeating out loud what I'd heard/read [REPEATING] while studying the image [FOCUSING ON IMAGES]. Sometimes they reminded me of other words/thing that would help me remember the new word. For example, the orange gause thing, a Fammel I think, reminded me of the word Flannel – the object had holes in it, an flannels are porous... a bit of an odd and tenuous connection, but that's what helped me to remember it. [ASSOCIATING WITH SOMETHING FAMILIAR]
English	5327303	associating the word with the picture [ASSOCIATING]
English	5328002	some sounded french and some sounded like english words [ASSOCIATING WITH SOMETHING FAMILIAR], and some I associated the first letter with something in the picture [ASSOCIATING WITH LETTER SHAPES]
English	5330132	i dont know i just did [*EXCL - NO STRATEGY*]
English	5332522	I repeated them out loud [REPEATING]
English	5332693	Repeating in my head what was said [REPEATING]
English	5333575	saying them in my head [REPEATING] and concentrating on the visual image and the colors in the image.[FOCUSING ON IMAGES]
English	5335169	By repeating the word in my head [REPEATING] while trying to remember the picture [FOCUSING ON IMAGES]
English	5336201	i paid particular attention to the beginning letter [FOCUSING ON LETTERS] as this is what was varying between the new words [COMPARING/CONTRASTING]. I also sometimes used the shape of

		the item eg macum looked a bit like an "m" so I used that to remember [ASSOCIATING WITH LETTER SHAPES]. But definitely the repetition [REPEATING] helped more than anything [EVALUATING LEARNING]
English	5337460	I hadn't had time to learn the spellings, so what I could remember was based on the sounds I could still (just about!) associate with the images [ASSOCIATING].
English	5338550	find something on the picture that 'sounds like' part of the word or imagine a friend with a similar name holding it. Eg for mackum i imagined 'smack um' because it looked like a kids toy you hit, or for madis i imagined my friend 'Madders' holding it [ASSOCIATING WITH SOMETHING FAMILIAR]
English	5338600	By associating the the picture with something that the name sounded like [ASSOCIATING WITH SOMETHING FAMILIAR]
English	5339584	sounds and colours compared to the other example [COMPARING/ CONTRASTING]
English	5339906	find something about the picture that i associated with part of the word [ASSOCIATING]
English	5340413	As explained above, I sometimes associated them with something that looks or sounds familiar (English word). [ASSOCIATING WITH SOMETHING FAMILIAR]
English	5341833	Notes with pictures [*EXCL - MAKING NOTES*]
English	5341870	By focusing on an aspect of the object that I could identify with English. For example, masim I learnt as it looked like a mass of small circular modules coming out of it [ASSOCIATING WITH SOMETHING FAMILIAR]. Mind you, it wasn't a very successful strategy. [EVALUATING LEARNING]
English	5343885	Combination of listening [FOCUSING ON SOUNDS] and reading [FOCUSING ON LETTERS]
English	5344305	Association to another word or image in my first language [ASSOCIATING WITH SOMETHING FAMILIAR]
English	5353998	related them to a ourpose [ASSOCIATING WITH SOMETHING FAMILIAR]

English	5354235	i pictured the initial letter in the picture [ASSOCIATING WITH LETTER SHAPES]
English	5356536	image association and sounds [ASSOCIATING]
English	5356537	looking at the shape to prompt / making up clues [FOCUSING ON IMAGES]
English	5357670	Visually [FOCUSING ON IMAGES] and through sound [FOCUSING ON SOUNDS]
English	5359677	Distinctive sounds [FOCUSING ON SOUNDS], especially first letter [FOCUSING ON LETTERS]
English	5366758	I tried to associate the picture with an english word that reminded me of that sound e.g. the black object looked mechanical so i thought of 'v' for vaccum, the flower object started with an 'f', the massid one looked like a mess! [ASSOCIATING WITH SOMETHING FAMILIAR]
English	5366866	randim no thought [*EXCL - NO STRATEGY*]
English	5374409	shape of object [FOCUSING ON IMAGES]
English	5374964	by associating the pictures with more familiar objects [ASSOCIATING WITH SOMETHING FAMILIAR]
English	5380819	comparing them to other familiar things [ASSOCIATING WITH SOMETHING FAMILIAR]
English	5384030	look at the shape and see if there was any link to the sound. ie the image for makum looked like an m shape [ASSOCIATING WITH LETTER SHAPES]
English	5384032	i tried to remember the voice saying it [FOCUSING ON SOUNDS]
English	5384657	Sound [FOCUSING ON SOUNDS]
English	5409941	sound of word [FOCUSING ON SOUND] and picture. [FOCUSING ON IMAGES]
English	5417849	Associated the images with the name, using the first letter to make a shape in my mind to help remember. [ASSOCIATING WITH LETTER SHAPES]

English	5418540	I repeated the word in my head [REPEATING] whilst linking it to the object being shown. [ASSOCIATING]
English	5434242	I said them out loud twice after each was shown, I then repeated the names once more before the testing phase [REPEATING]. I almost had it in a rhythm that kept the learning time very fast-paced but consistent [DIRECTING LEARNING]. I did, however, think that the learning phase was the timed game and so was intentionally learning them quickly rather than learning them well [EVALUATING LEARNING].
English	5499381	just from the sound and used the game like a memory game with pictures [FOCUSING ON SOUNDS]
English	5499420	I put them in a sentence or described what they vaguely looked like. For example, "a nakem is a flip flop" [ASSOCIATING WITH SOMETHING FAMILIAR], "a nasit has 5 holes" [ASSOCIATING WITH VISUAL DETAIL]
English	5499421	Toward the end of the learning phase, when each of the four objects were shown, I tried to name each one (rather than just the one I was asked to identify) [DIRECTING LEARNING]
English	5502013	relate the pictures with imagery and initial letter sounds [ASSOCIATING]
English	5503474	I mouthed or said them aloud as I went through [REPEATING]. I also sometimes made a mental note of a distinctive feature of the picture, like 'blue blob' or 'upside-down yellow Y' [ASSOCIATING WITH VISUAL DETAIL]
English	5510526	I just tried to remember the pictures that I saw and at least one of the words [REMEMBERING]. That way when I heard the word I remembered, I clicked the right picture and I knew the other would be the other picture I'd just seen. Sometimes I remembered them both! [DIRECTING LEARNING]
English	5521508	I sometimes associated the shape of the object with femimine sounds [m] and [f]. 'Madis' had a curvy shape which I associated with [m] as in mother. I also associated similar sounding English words or phrases to the way the objects looked, e.g. the 'vamil' had humps like a 'camel'. [ASSOCIATING WITH SOMETHING FAMILIAR]

English	5557001	Associating the sounds of the words with characterstics of the pictures. 'Madis' looked like a hippo, so I decided it was a mad hippo. 'Famil' looked like a 'family' of spheres etc. [ASSOCIATING WITH SOMETHING FAMILIAR]
English	5586643	colour association [ASSOCIATING WITH VISUAL DETAIL]
English	5586843	colours to sounds - so if it said famas (sorry If I have that one wrong) i remembered orange and looks like plastic beads strung together [ASSOCIATING WITH VISUAL DETAIL]
English	5587514	studying the picture [FOCUSING ON IMAGES] whilst repeating the word over and over [REPEATING]
English	5588274	i used strategies like looks like/sounds like. eg, famis - looks like its proud of itself for being famous. nadit - looks like a rocket and sounds like a rocket. nackum - looks like something to whack with and sounds like 'whack 'em'. mackum - looks like an alien which i named Mack. didnt have them for all 12 but it helped me remember the difference [ASSOCIATING WITH SOMETHING FAMILIAR]
English	5592006	said them out load to myself while looking at the picture [FOCUSING ON IMAGE], and repeated until image disappeared [REPEATING]
English	5595978	There are some pictures that I tried to match with something like there was a vedet? which was a picture of something that looked like an upside down V to me? [ASSOCIATING WITH LETTER SHAPES] And then the picture that looked like a sandal was near the beginning and I thought it sounded like nechem...that one seemed to stick in my brain easier, I don't know why? [ASSOCIATING WITH SOMETHING FAMILIAR]
English	5602966	I sometimes tried to associate a letter with what I was seeing, e.g. was there a 'v' shape in any of the pictures that had a 'v' in the word? [ASSOCIATING WITH LETTER SHAPES] I also tried to assign some meaning to the object I was seeing which were all random objects to me. E.g. with the word famil (if I've rememebr that right), it looked a bit like an object that could create a funnel if inserted somewhere! [ASSOCIATING WITH SOMETHING FAMILIAR]
English	5615580	i didn't 'try' anything I wouldn't normally just naturally do [*EXCL - NO STRATEGY*]

English	5620967	Spoke them out loud, repeated the words in my head [REPEATING]. Focused on the first letter [FOCUSING ON FIRST LETTER] and sound [FOCUSING ON SOUND].
English	5621683	Associating the word with something about the picture. [ASSOCIATING]
English	5627982	remembering the image [FOCUSING ON IMAGES]
English	5627983	sounds recognition [FOCUSING ON SOUNDS]
English	5629029	something in the spelling or sound to link to the picture [ASSOCIATING]
English	5637584	picture association [ASSOCIATING]
English	5642980	Rembering the first letter of the word [FOCUSING ON LETTERS] and placing it with the colour of the object in my memory [ASSOCIATING WITH VISUAL DETAIL]
English	5649032	By making connections [ASSOCIATING] between the sound of the word (more than its spelling - but thinking about what letters I thought I was hearing in it [FOCUSING ON SOUNDS]) and the image of the object the word described.
English	5655323	Repetition [REPEATING]
English	5659195	i focused on the first letter [FOCUSING ON LETTERS] as some words sounded similar apart from the first letter [COMPARING/CONTRASTING]
English	5659326	I repeated them to myself out loud after hearing and or seeing them. [REPEATING]
English	5663887	analogies of things they reminded me of. e.g. the famel made me think of camel humps, so in my mind it was a Fat cAMEL [ASSOCIATING WITH SOMETHING FAMILIAR]
English	5664502	creating a blank in my mind like "_ A _ I _ " where consonants would go in the three blanks [VISUALISING] because all examples were CVCVC [COMPARING/CONTRASTING], and associating with the traits of the images (like how many prongs they had toward the left or bottom, or what colors they were) [ASSOCIATING WITH VISUAL DETAIL]

English	5666899	Repeated the word in my head [REPEATING]
English	5667662	I used the sound of the word and the shape of the object to create a playful image in my mind [VISUALISING]
English	5670727	the sounds I remember from the pronunciation which i associated with the picture. [ASSOCIATING]
English	5671003	I said them aloud after they were said [REPEATING]
English	5672324	Pneumonic device [REMEMBERING]
English	5674519	associated with an english word that the picture made me think of [ASSOCIATING WITH SOMETHING FAMILIAR]
English	5676564	I tried to visually see if something in the image could remind me of the start of tbe letter e.g. famil having circular orange parts I linked the "amil" of the word [ASSOCIATING WITH LETTER SHAPES] I also just tried to use photographic memory to take a mental picture of the word and picture [VISUALISING]
English	5690497	I repeated them in my head when the picture was shown [REPEATING]. I also tried to link the words to the images in my head. [ASSOCIATING]
English	5691858	Not through the spelling apparently.. Through the visuals of the object [FOCUSING ON IMAGES] and the repitition of the sounds. [REPEATING]
English	5692078	Phonetically [*EXCL - INSUFFICIENT*]
English	5692100	Said them out loud [REPEATING]
English	5696781	Linking shapes for the letters with the object – eg Mackem had 3 legs, like an M [ASSOCIATING WITH LETTER SHAPES]
English	5704997	I picked out a part of the image that i felt related to the word (the words with a V at the start i looked for a V shape in the object) [ASSOCIATING WITH LETTER SHAPES]
English	5706490	repeat out loud and internally [REPEATING]
English	5706491	Through creating shapes of the letters within the pictures of the letter that the word began with [ASSOCIATING WITH LETTER SHAPES] OR by creating memories to try and relate them too. For example, the hippo

		looking toy reminded me of Gloria from Madagascar and the letter the word began with was an M so when it came to that word I was able to relate one to another. [ASSOCIATING WITH SOMETHING FAMILIAR]
English	5719579	Look at the shapes/colors of the objects [FOCUSING ON IMAGES] and say the words to myself [REPEATING]. If spelling included, in English, I included that with my memory of the object. [VISUALISING]
English	5732647	Tried to link the new word image to a more familiar object or concept [ASSOCIATING WITH SOMETHING FAMILIAR]
English	5750023	I tried to associate the picture with the sound (i.e. Nasset looked like a vase with holes in it, nass in German means wet- the floor would be wet if you put water in that vase, Makham looked a bit like the Tyneside bridge in Newcastle and that's what they call people from Sunderland etc...) [ASSOCIATING WITH SOMETHING FAMILIAR]
English	5750994	Repeated them in my head multiple times [REPEATING], visualizing how the words sound incoproated with how they're spelt [VISUALISING]
English	5751031	focusing on the consonants, especially those at the beginning of each syllable [FOCUSING ON SOUNDS]. I also realized you were mostly using words with minimal pairs (f/v, m/n) so focused on those especially. [COMPARING/ CONTRASTING]
English	5755297	By coming up with acronyms with the first letter [FOCUSING ON LETTERS] in the word to associate it with the image. [ASSOCIATING]
English	5755298	association of the sounds with the image displayed [ASSOCIATING] by repeating the word and pointing at the image [REPEATING]
English	5755299	memorizing them as a chant with distinguishing colors and shapes [ASSOCIATING WITH VISUAL DETAIL]
English	5755300	In earlier sections, by the first syllable [FOCUSING ON SOUNDS], then when two words had the same syllable I imagined that the second syllable was related to what the object does (even if it might not be what the object is meant for in reality) [ASSOCIATING WITH SOMETHING FAMILIAR]
English	5755301	I sounded out the words as I heard them and saw the pictures [REPEATING]

English	5756084	I focused on the contrasts in sounds [FOCUSING ON SOUNDS]
English	5756085	Graft sound onto object [ASSOCIATING]
English	5758767	I thought about what the pictures looked like and tried to create associations even if they were weird. eg I think it was 'famil' that I remembered as lots of little loops making a family. Or for 'nackam' I thought about a back being 'knackered' and that it looked a bit like someone lying on their back. [ASSOCIATING WITH SOMETHING FAMILIAR]
English	5759398	matching shape to letters [ASSOCIATING WITH LETTER SHAPES]
English	5759735	Sounds [FOCUSING ON SOUNDS], shapes [FOCUSING ON IMAGES]
English	5762777	tried relating the pictures to known objects and using mnemonics, ex: masit looked like a mass of something [ASSOCIATING WITH SOMETHING FAMILIAR]
English	5765135	i tried to memorise which words had which sounds [FOCUSING ON SOUNDS] and then I attempted to memorise the first letter [FOCUSING ON LETTERS] which was often the difference between the two similar sounds eg the voiced vammell and voiceless fammell [COMPARING/CONTRASTING]
English	5765136	By trying to find a word in English that sounds similar or has similar sounds in it and also by trying to find a similarity between the sound and the object, for example the "fammal" reminded me a bit of a fan, so it was easier to remember it. [ASSOCIATING WITH SOMETHING FAMILIAR]
English	5780967	map on to a similar english word or do an action [ASSOCIATING WITH SOMETHING FAMILIAR]
English	5831108	try to associate the made up words with what i thpught it kind of looked like [ASSOCIATING WITH SOMETHING FAMILIAR]
English	5884960	Sound [FOCUSING ON SOUNDS]
English	5909329	associate the sound with the picture [ASSOCIATING]
English	5909763	Initially I tried associating the picture with a colour or a description of the item [ASSOCIATING WITH VISUAL DETAIL]. I then tried some

		repetition of the word [REPEATING]
English	5914367	Repeating the word in my head [REPEATING]
English	5998960	I found 'Mackem' the easiest to remember, as this is a real English word for someone from (or the dialect of) Sunderland / Wearside [ASSOCIATING WITH SOMETHING FAMILIAR]. For everything else, I tried to associate the word with the colour of the object rather than the object itself as a shortcut [ASSOCIATING WITH VISUAL DETAIL], but this got hard when there were three orange objects. [EVALUATING LEARNING]
English	6029761	I remembered the first letter [FOCUSING ON LETTERS] or sound [FOCUSING ON SOUNDS] of each word and tried to associate it with something to do with the picture [ASSOCIATING]

Appendix XIX: LLS data pre-processing Rmd output

This document outlines the steps to process responses from the debrief questionnaire, including the extraction of responses relevant to participant language learning strategies (LLS). The structure of the document is as follows:

1. Data cleaning and exclusions
2. Reliability of strategy inventory

Data cleaning and exclusions

Raw data from the demographic questionnaire, post-test questionnaire and matching task are initially cleaned and reformatted into excel, then each read into R. Data types are then corrected where necessary, followed by adjusting labels and levels of ordered variables.

The variables relevant to analysis of LLS are selected from the post-test questionnaire and merged with the relevant demographic and task performance variables in a new data set.

Exclusions are then calculated based on the same criteria as the word learning and matching data. Specifically participants are excluded if they:

- were under 16 or over 65 years of age.
- reported problems with vision or hearing.
- reported a disproportionate amount of distraction during participation.
- were English and reported having literacy experience in Arabic or a language that uses the Arabic script.
- did not complete the word learning and matching task on the same day.

After exclusions, 114 L1 Arabic and 117 L1 English participants remain.

Reliability of strategy inventory

The next step is to inspect the strategy inventory items and recode items numerically (1="never", 2 ="almost never", 3="sometimes", 4="almost always", 5="always").

The internal reliability of the inventory is then calculated using Cronbach's alpha. The wording is positive for all items, apart from the statements about ignoring written forms. So, reverse coding was used for those two items. The R script below calculated an acceptable reliability of α = 0.75 for the full inventory. Reliability of subscales within the inventory are calculated as part of the PCA analysis and available in appendix XX.

```
#recode likert scales
likert_recode <- function(x){
  as.numeric(case_when(
```

```
    x == "never" ~ 1,
    x == "almost_never" ~ 2,
    x == "sometimes" ~ 3,
    x == "almost_always" ~ 4,
    x == "always" ~ 5,
  ))
}

likert_recode_negative <- function(x){
  as.numeric(case_when(
    x == "never" ~ 5,
    x == "almost_never" ~ 4,
    x == "sometimes" ~ 3,
    x == "almost_always" ~ 2,
    x == "always" ~ 1,
  ))
}

statements_positive <- strategy_rename %>%
  select(action, associations, context, mental_image, sounds_letters,
sublexical, first_sound, distinguish_arabic, distinguish_english,
similarities, grouping, patterns, repeating, mouthing,
visualise_english, visualise_arabic, mistakes, progress, recalling,
relaxed_english, relaxed_arabic)%>%
  mutate_all(likert_recode)

statements_negative <- strategy_rename %>%
  select(ignore_arabic, ignore_english) %>%
  mutate_all(likert_recode_negative)

strategy_recode <- cbind(statements_positive, statements_negative)

alpha(strategy_recode)
```

Appendix XX: LLS PCA Rmd output

This document outlines the procedure and provides the R script for conducting a principal components analysis with the language learning strategy (LLS) inventory data for both language groups (psych package). The factor scores from the L1 Arabic group are then further explored using correlation and regression analyses (appendix XXI). Relevant script and output are included below.

After removing the missing data for the 5 L1 Arabic participants who did not give answers for any of the inventory items, new data frames are created selecting only the strategy inventory items for each language group.

Before conducting the PCA analysis, checks are conducted to look for variables that:
1. don't correlate with other variables (look out for few correlations r > .3)
2. correlate very highly (r = .9)
3. check Bartlett's test of sphericity is significant (p<.05)
4. check KMO statistic is > .5 as a bare minimum
5. check determinant of this matrix is bigger than 0.00001 (so multicollinearity isn't an issue)

L1 English PCA

First, it is important to assess whether particular items need to be excluded before continuing with the PCA. The outlined checks were conducted using the script below.

```
strategyEnMatrix <- cor(strategy_EnQs)

#eyeball correlations
round(strategyEnMatrix,2)
# "ignore_arabic" doesn't have any correlations >.3 - consider excluding

#check it's not an identity matrix - appears to be significant
cortest.bartlett(strategyEnMatrix, n=117)

KMO(strategyEnMatrix)
# KMO = .63 which is mediocre, "ignore_arabic" is .45 and "sublexical" is .50,
so grounds for exclusion

## rerun checks without "ignore arabic" and "sublexical" variable
strategy_EnQs2 <- strategy_EnQs %>%
  select(-c("ignore_arabic", "sublexical"))

strategyEnMatrix2 <- cor(strategy_EnQs2)
corrplot(strategyEnMatrix2,order="hclust", type="upper", tl.col="black",
tl.srt=45)
```

```
#eyeball correlations
round(strategyEnMatrix2,2)
# all have at least 1 correlation >.3

#check it's not an identity matrix - appears to be significant
cortest.bartlett(strategyEnMatrix2, n=117)

KMO(strategyEnMatrix2)
# KMO = .66 which is mediocre but marginally improved, all are above .5

det(strategyEnMatrix2)
# determinant is greater than necessary value, so not problematic
```

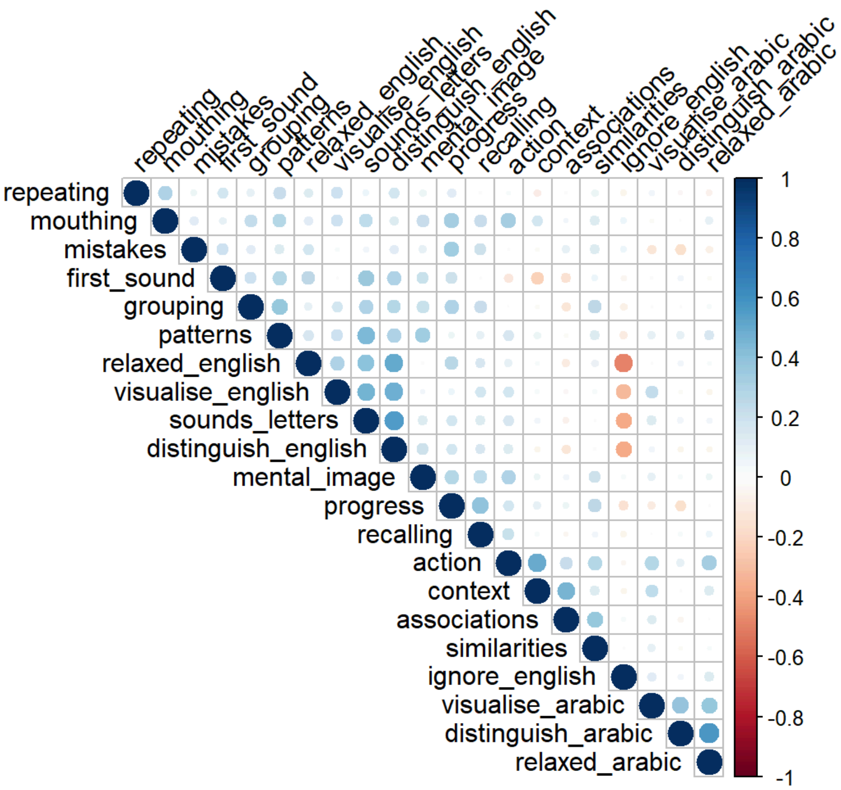


Figure XX.1: Correlation plot of all strategy items included in the L1 English PCA analysis

After making the necessary changes, such as excluding the strategies of *ignoring Arabic OI* and *breaking words down into smaller units*, the preliminary analysis is satisfactory for the L1 English data. To do this, I use the principal() function and extract the same number of factors as there are variables. Then, I inspect their eigenvalues and make decisions about factor extraction.

```
#PCA model 1 with same number of factors as variables

pc1_en <-principal(strategyEnMatrix2, nfactors =
length(strategyEnMatrix2[,1]), rotate = "none")
```

```
pc1_en
# 7 factors have eigenvalues greater than 1 (Kaiser's criterion)

plot(pc1_en$values, type = "b")
# this demonstrates evidence of a plateau after 5 factors...but
subjective

# additional analysis to consider how many factors to extract
results_En <- nScree(x=pc1_en$value)
results_En
plotnScree(results_En)
```

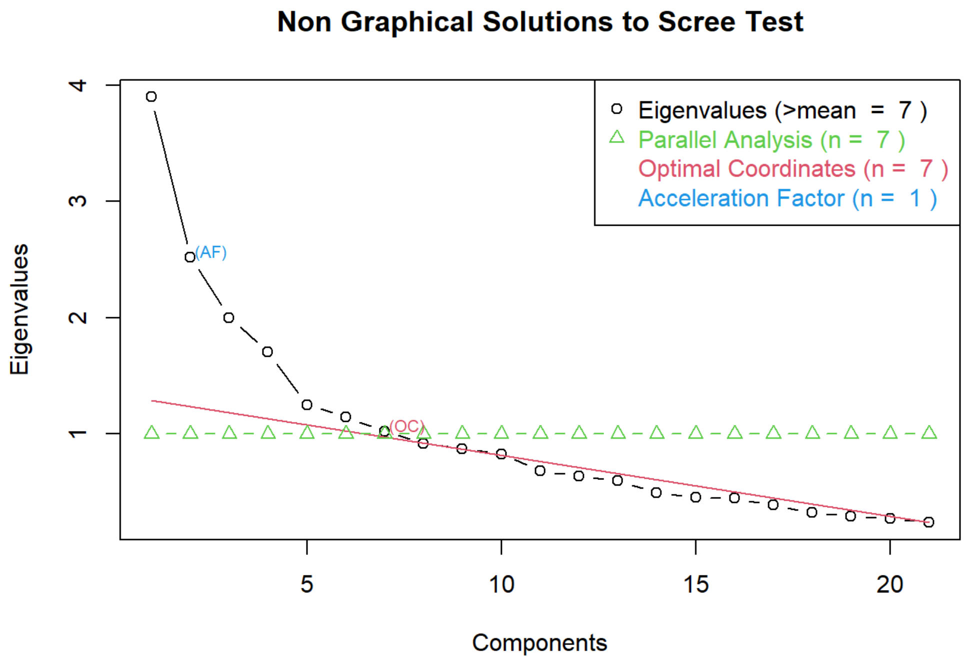


Figure XX.2: Scree plot including Parallel Analysis and Optimal Coordinates to assess number of factors to extract for the L1 English data

Kaiser's criterion of including factors with eigenvalues over 1, as well as looking at the scree plot through the lens of parallel analysis and optimal coordinates, provides grounds to continue with PCA extracting 7 factors. This decision is additionally assessed by examining communalities of variables and residuals.

```
# continue with analysis extracting 7 factors
pc2_eng <-principal(strategyEnMatrix2, nfactors = 7, rotate = "none")
pc2_eng

mean(pc2_eng$communality)
```

```
# commonality (h2 column) average is not greater than .7, which does not
meet Kaiser's criterion (=.64) - worth re-running analysis with
additional factors. Also, fit = .88 - not great, as .90 is a good rule
of thumb as a minimum

# we can further check this by looking at residuals
residuals_eng2 <-factor.residuals(strategyEnMatrix2, pc2_eng$loadings)
residuals_eng2 <-as.matrix(residuals_eng2[upper.tri(residuals_eng2)])
large.resid_eng2 <-abs(residuals_eng2)>0.05

sum(large.resid_eng2)
sum(large.resid_eng2)/nrow(residuals_eng2)
#less than 50% of residuals greater than 0.05, so this is ok - but .45
is not ideal

sqrt(mean(residuals_eng2^2))
#quite high but not awful

hist(residuals_eng2)
#approximately normal but slight negative skew
```

Based on these assessments, I rerun the analysis with 8 factors to check improved fit.

```
# continue with analysis extracting 8 factors
pc3_eng <-principal(strategyEnMatrix2, nfactors = 8, rotate = "none")
pc3_eng

mean(pc3_eng$communality)

#commonality (h2 column) average is not greater than .7, which does not
meet Kaiser's criterion (=.69) - worth re-running analysis with more
factors. Fit = .89 - over .90 is a good rule of thumb as a min

#we can further check this by looking at residuals
residuals_eng3 <-factor.residuals(strategyEnMatrix2, pc3_eng$loadings)
residuals_eng3 <-as.matrix(residuals_eng3[upper.tri(residuals_eng3)])
large.resid_eng3 <-abs(residuals_eng3)>0.05

sum(large.resid_eng3)
sum(large.resid_eng3)/nrow(residuals_eng3)
#less than 50% of residuals greater than 0.05, so .38 is a little better

sqrt(mean(residuals_eng3^2))
#quite high but better
```

```
hist(residuals_eng3)
#slight skew but improved
```

Based on these assessments, I rerun the analysis with 9 factors to check improved fit.

```
# continue with analysis extracting 8 factors
pc4_eng <-principal(strategyEnMatrix2, nfactors = 9, rotate = "none")
pc4_eng

mean(pc4_eng$communality)

#commonality (h2 column) for most variables is greater than .7 and
average =.73, meeting Kaiser's criterion. Also, fit = .91

#we can further check this by looking at residuals
residuals_eng4 <-factor.residuals(strategyEnMatrix2, pc4_eng$loadings)
residuals_eng4 <-as.matrix(residuals_eng4[upper.tri(residuals_eng4)])
large.resid_eng4 <-abs(residuals_eng4)>0.05

sum(large.resid_eng4)
sum(large.resid_eng4)/nrow(residuals_eng4)
#less than 50% of residuals greater than 0.05, so .33 is better

sqrt(mean(residuals_eng4^2))
# better

hist(residuals_eng4)
# slight skew
```

Once satisfied with the appropriate checks, the PCA continues with the extraction of 9 factors. To improve interpretation of factors, rotation maximises the loading of each variable on the extracted factors. As it is assumed that underlying dimensions are likely to be related to each other, an oblique rotation is applied to the data.

```
# oblique rotation

pc6_eng <- principal(strategyEnMatrix2, nfactors = 9, rotate =
"oblimin")
pc6_eng

print.psych(pc6_eng, cut = 0.3, sort = TRUE)
```

```
## Principal Components Analysis
## Call: principal(r = strategyEnMatrix2, nfactors = 9, rotate = "oblimin")
## Standardized loadings (pattern matrix) based upon correlation matrix
##                     item   TC2   TC4   TC1   TC3   TC5   TC8   TC7   TC9   TC6
## context                3  0.84
## action                 1  0.72
## associations           2  0.59                                          0.40
## patterns              11        0.74
## mental_image           4        0.66
## first_sound            6        0.60                                          0.37
## sounds_letters         5        0.54  0.40              0.31
## ignore_english        16             -0.85
## relaxed_english       20              0.80
## distinguish_english    8        0.39  0.42              0.39
## distinguish_arabic     7                    0.88
## relaxed_arabic        21                    0.84
## recalling             19                          0.84
## progress              18                          0.57
## visualise_english     14                                0.81
## visualise_arabic      15                    0.51        0.57
## repeating             12                                      0.88
## mouthing              13                          0.31        0.66
## similarities           9                                            0.86
## grouping              10 -0.34  0.34              0.32              0.55
## mistakes              17                                                  0.78
##                       h2   u2 com
## context             0.71 0.29 1.1
## action              0.70 0.30 1.5
## associations        0.71 0.29 2.9
## patterns            0.64 0.36 1.2
## mental_image        0.58 0.42 1.9
## first_sound         0.73 0.27 3.1
## sounds_letters      0.70 0.30 2.6
## ignore_english      0.76 0.24 1.2
## relaxed_english     0.75 0.25 1.3
## distinguish_english 0.68 0.32 3.4
## distinguish_arabic  0.77 0.23 1.1
## relaxed_arabic      0.76 0.24 1.1
## recalling           0.72 0.28 1.2
## progress            0.70 0.30 3.0
## visualise_english   0.82 0.18 1.3
## visualise_arabic    0.75 0.25 2.8
## repeating           0.81 0.19 1.2
## mouthing            0.72 0.28 2.2
## similarities        0.80 0.20 1.1
```

```
## grouping            0.77 0.23 3.9
## mistakes            0.74 0.26 1.4
##
##                       TC2  TC4  TC1  TC3  TC5  TC8  TC7  TC9  TC6
## SS loadings           2.04 2.14 2.07 1.92 1.69 1.51 1.41 1.39 1.16
## Proportion Var        0.10 0.10 0.10 0.09 0.08 0.07 0.07 0.07 0.06
## Cumulative Var        0.10 0.20 0.30 0.39 0.47 0.54 0.61 0.68 0.73
## Proportion Explained  0.13 0.14 0.14 0.13 0.11 0.10 0.09 0.09 0.08
## Cumulative Proportion 0.13 0.27 0.41 0.53 0.64 0.74 0.83 0.92 1.00
##
##  With component correlations of
##       TC2   TC4   TC1   TC3   TC5   TC8   TC7   TC9   TC6
## TC2  1.00 -0.06 -0.06  0.11  0.07  0.07  0.01  0.15  0.01
## TC4 -0.06  1.00  0.20  0.09  0.21  0.13  0.21  0.14  0.05
## TC1 -0.06  0.20  1.00 -0.09  0.10  0.19  0.11  0.00  0.05
## TC3  0.11  0.09 -0.09  1.00 -0.02  0.09 -0.02 -0.01 -0.03
## TC5  0.07  0.21  0.10 -0.02  1.00  0.02  0.13  0.10  0.02
## TC8  0.07  0.13  0.19  0.09  0.02  1.00  0.08  0.03 -0.07
## TC7  0.01  0.21  0.11 -0.02  0.13  0.08  1.00  0.10  0.08
## TC9  0.15  0.14  0.00 -0.01  0.10  0.03  0.10  1.00  0.10
## TC6  0.01  0.05  0.05 -0.03  0.02 -0.07  0.08  0.10  1.00
##
## Mean item complexity =  1.9
## Test of the hypothesis that 9 components are sufficient.
##
## The root mean square of the residuals (RMSR) is  0.06
##
## Fit based upon off diagonal values = 0.91
```

```
# report the structure matrix as well as the pattern matrix above (but a
little hard to make sense of) - look for loadings >.4 for both matrices.

pc6_engStr<- pc6_eng$loadings %*% pc6_eng$Phi
pc6_engStr

# Noteworthy that correlations between subscales are low, where there is
just a small correlation between TC1 and TC4, which is logical

# Also important to note that the four subscales which are extracted only
explain 39% of the variance, below the 50% recommended threshold (Streiner,
1994)
```

```
##                             TC2          TC4          TC1          TC3          TC5
## action               0.74396724   0.18142563   0.05113096   0.28094574   0.29647274
```

```
## associations         0.63566899 -0.15990048 -0.11814234 -0.05757072 -0.20406053
## context              0.83485195 -0.01604159  0.01260175  0.07267192  0.03242095
## mental_image         0.21891545  0.62892593 -0.11929746  0.05368264  0.35079145
## sounds_letters       0.01793048  0.63239744  0.55505234  0.05031553  0.13333538
## first_sound         -0.32780895  0.62471212  0.27461279  0.07894110 -0.07392409
## distinguish_arabic  -0.03021750  0.06745721 -0.02538593  0.85989429 -0.06198713
## distinguish_english -0.07938021  0.52148692  0.59576278 -0.08302488  0.17417666
## similarities         0.24997942  0.09104017  0.01921711  0.02911714  0.05917618
## grouping            -0.24259860  0.49807887  0.11078038 -0.01096688  0.42597785
## patterns             0.04654269  0.76206930  0.10951176  0.12952132  0.08094898
## repeating           -0.09599052  0.11759419  0.08862049 -0.03981911 -0.08466131
## mouthing             0.28114958  0.31712703  0.03784053  0.02896816  0.42699167
## visualise_english    0.04330364  0.17346689  0.35443648 -0.06898762  0.19432164
## visualise_arabic     0.27139362  0.07506802 -0.17340828  0.59664258 -0.12648519
## ignore_english      -0.01641878 -0.07663176 -0.82208271  0.16467682 -0.03976587
## mistakes            -0.03197988  0.14380368  0.05189356 -0.16768965  0.29364297
## progress             0.12407372  0.20644779  0.30277700 -0.10483625  0.64223146
## recalling            0.06600092  0.14031661  0.07970916  0.05290795  0.82322341
## relaxed_english     -0.01446615  0.20500619  0.81710066  0.08035696  0.15126470
## relaxed_arabic       0.21398912  0.09876284 -0.08238110  0.84767780  0.11599720
##                              TC8         TC7         TC9         TC6
## action               0.176942462  0.16104266  0.13227634 -0.05653965
## associations         0.030189289 -0.05559968  0.47060476  0.26504843
## context              0.080788752 -0.03218118  0.16026966 -0.06811194
## mental_image        -0.019232791  0.08972425  0.19443903  0.07991983
## sounds_letters       0.463119542  0.14192604  0.02218258 -0.01549331
## first_sound         -0.118721996  0.19926370  0.06104137  0.41789678
## distinguish_arabic   0.076018688 -0.05045007 -0.01417401 -0.09757276
## distinguish_english  0.501563950  0.18786097 -0.01092112  0.11058555
## similarities         0.003300829  0.12655923  0.87745835  0.16178860
## grouping             0.154391293  0.15933350  0.55722580 -0.21886222
## patterns             0.197057156  0.30067281  0.19643749 -0.03870539
## repeating            0.162723424  0.86324382  0.10570466  0.11285587
## mouthing            -0.004712068  0.70605261  0.09081387 -0.06176452
## visualise_english    0.846441154  0.24400412  0.02543354 -0.07478617
## visualise_arabic     0.573102315 -0.03091936  0.12660852  0.06368327
## ignore_english      -0.192632309  0.01479356 -0.03809837  0.17777810
## mistakes            -0.035507196  0.13086629  0.17219344  0.79049462
## progress            -0.161750514  0.27448914  0.35525215  0.32500992
## recalling            0.156459617  0.03342864  0.06308858  0.11582966
## relaxed_english      0.188938158  0.19294443  0.01494862  0.26033767
## relaxed_arabic      -0.007006466  0.01294324 -0.01167543 -0.05393853
```

Next, it is necessary to conduct reliability analysis on subscales where at least three items have a loading >.4 and do not load highly across multiple other factors. There are four subscales that meet this criteria:

TC2 = Connecting with meaning

TC4 = Analysing sounds and letters
TC1 = Using English OI
TC3 = Using Arabic OI

```
connectingMeaning <- strategy_EnQs2 [, c("context", "action",
"associations")]

analysingSoundsLetters <- strategy_EnQs2 [, c("sounds_letters",
"patterns", "mental_image", "first_sound")]

usingEngOI <- strategy_EnQs2[, c("ignore_english", "relaxed_english",
"distinguish_english", "sounds_letters")]

usingArOI <- strategy_EnQs2 [, c("distinguish_arabic", "relaxed_arabic",
"visualise_arabic")]
```

Calculating the alpha is straightforward for all, except for the third subscale *using English OI*. This contains "ignore_english" which needs to be negatively coded.

```
alpha(connectingMeaning)
# overall reliability is acceptable (alpha =.65), none of the items
would increase reliability if they were deleted and the r.drop are all
above .3.
```

```
##
## Reliability analysis
## Call: alpha(x = connectingMeaning)
##
##   raw_alpha std.alpha G6(smc) average_r S/N   ase mean   sd median_r
##       0.65      0.66     0.6      0.39 1.9 0.055  2.2 0.86     0.46
##
##     95% confidence boundaries
##          lower alpha upper
## Feldt     0.52  0.65  0.75
## Duhachek  0.54  0.65  0.76
##
##  Reliability if an item is dropped:
##              raw_alpha std.alpha G6(smc) average_r  S/N alpha se var.r med.r
## context           0.35      0.37    0.23      0.23 0.58    0.112    NA  0.23
## action            0.62      0.63    0.46      0.46 1.68    0.069    NA  0.46
## associations      0.66      0.67    0.50      0.50 2.00    0.062    NA  0.50
##
##  Item statistics
##                  n raw.r std.r r.cor r.drop mean  sd
```

```
## context      117  0.84  0.84  0.74   0.60  1.8 1.1
## action       117  0.68  0.75  0.55   0.41  1.5 0.9
## associations 117  0.78  0.73  0.50   0.41  3.1 1.3
##
## Non missing response frequency for each item
##                 1    2    3    4    5 miss
## context      0.56 0.21 0.14 0.07 0.03    0
## action       0.69 0.14 0.12 0.05 0.00    0
## associations 0.16 0.14 0.26 0.29 0.15    0
```

```
alpha(analysingSoundsLetters)
# overall reliability is acceptable (alpha =.63), none of the items
would increase reliability if they were deleted and the r.drop values
are above .3. However, "mental_image" = .30 is low.
```

```
##
## Reliability analysis
## Call: alpha(x = analysingSoundsLetters)
##
##   raw_alpha std.alpha G6(smc) average_r S/N   ase mean   sd median_r
##       0.63      0.63    0.58       0.3 1.7 0.055  3.4 0.87     0.31
##
##     95% confidence boundaries
##          lower alpha upper
## Feldt     0.51  0.63  0.73
## Duhachek  0.52  0.63  0.74
##
##  Reliability if an item is dropped:
##                  raw_alpha std.alpha G6(smc) average_r  S/N alpha se  var.r
med.r
## sounds_letters      0.54      0.54    0.44      0.28 1.16    0.074 0.0042
0.28
## patterns            0.49      0.48    0.40      0.24 0.94    0.080 0.0139
0.21
## mental_image        0.63      0.63    0.54      0.36 1.72    0.059 0.0061
0.37
## first_sound         0.58      0.57    0.50      0.31 1.32    0.067 0.0244
0.34
##
##  Item statistics
##                  n raw.r std.r r.cor r.drop mean  sd
## sounds_letters 117  0.72  0.71  0.57   0.45  3.1 1.3
## patterns       117  0.76  0.75  0.64   0.51  3.1 1.3
## mental_image   117  0.58  0.61  0.39   0.30  3.6 1.1
## first_sound    117  0.69  0.68  0.49   0.39  3.7 1.3
##
## Non missing response frequency for each item
```

```
##                    1    2    3    4    5 miss
## sounds_letters 0.16 0.12 0.32 0.22 0.17    0
## patterns       0.18 0.15 0.25 0.27 0.15    0
## mental_image   0.07 0.08 0.26 0.38 0.21    0
## first_sound    0.11 0.05 0.21 0.28 0.34    0
```

```
alpha(usingEngOI, keys = c(-1, 1, 1, 1))
# overall reliability is good (alpha = .76), none of the items would increase reliability if they were deleted and the r.drop values are all above .3 which indicates a good amount of correlation
```

```
##
## Reliability analysis
## Call: alpha(x = usingEngOI, keys = c(-1, 1, 1, 1))
##
##   raw_alpha std.alpha G6(smc) average_r S/N   ase mean   sd median_r
##       0.76      0.77    0.72      0.45 3.3 0.035  3.5 0.98     0.44
##
##     95% confidence boundaries
##          lower alpha upper
## Feldt     0.69  0.76  0.83
## Duhachek  0.70  0.76  0.83
##
##  Reliability if an item is dropped:
##                     raw_alpha std.alpha G6(smc) average_r S/N alpha se  var.r
## ignore_english-          0.74      0.74    0.67      0.49 2.9    0.041 0.0061
## relaxed_english          0.70      0.70    0.62      0.43 2.3    0.046 0.0118
## distinguish_english      0.68      0.68    0.60      0.42 2.2    0.051 0.0032
## sounds_letters           0.71      0.71    0.63      0.45 2.5    0.045 0.0052
##                     med.r
## ignore_english-      0.50
## relaxed_english      0.37
## distinguish_english  0.40
## sounds_letters       0.48
##
##  Item statistics
##                       n raw.r std.r r.cor r.drop mean  sd
## ignore_english-     117  0.69  0.73  0.58   0.50  4.3 1.0
## relaxed_english     117  0.79  0.78  0.67   0.58  3.3 1.4
## distinguish_english 117  0.81  0.79  0.71   0.62  3.2 1.4
## sounds_letters      117  0.77  0.76  0.65   0.56  3.1 1.3
##
## Non missing response frequency for each item
##                         1    2    3    4    5 miss
## ignore_english      0.64 0.12 0.18 0.03 0.03    0
## relaxed_english     0.15 0.13 0.23 0.24 0.25    0
## distinguish_english 0.16 0.12 0.28 0.21 0.22    0
```

```
## sounds_letters      0.16 0.12 0.32 0.22 0.17    0
```

```
alpha(usingArOI)
# overall reliability is acceptable (alpha = .67). none of the items
would increase reliability if they were deleted, apart from perhaps
“visualise_arabic”, and the r.drop values are all above .3.
```

```
## 
## Reliability analysis   
## Call: alpha(x = usingArOI)
## 
##   raw_alpha std.alpha G6(smc) average_r S/N   ase mean   sd median_r
##       0.67      0.71    0.63      0.44 2.4 0.051  1.1 0.36     0.39
## 
##     95% confidence boundaries 
##          lower alpha upper
## Feldt     0.56  0.67  0.76
## Duhachek  0.57  0.67  0.77
## 
##  Reliability if an item is dropped:
##                   raw_alpha std.alpha G6(smc) average_r S/N alpha se var.r
## distinguish_arabic      0.54      0.54    0.37      0.37 1.2    0.084    NA
## relaxed_arabic          0.52      0.56    0.39      0.39 1.3    0.079    NA
## visualise_arabic        0.68      0.73    0.57      0.57 2.7    0.051    NA
##                   med.r
## distinguish_arabic  0.37
## relaxed_arabic      0.39
## visualise_arabic    0.57
## 
##  Item statistics 
##                     n raw.r std.r r.cor r.drop mean   sd
## distinguish_arabic 117  0.76  0.82  0.70   0.58  1.1 0.33
## relaxed_arabic     117  0.83  0.82  0.69   0.53  1.2 0.53
## visualise_arabic   117  0.77  0.74  0.50   0.42  1.2 0.52
## 
## Non missing response frequency for each item
##                       1    2    3    4 miss
## distinguish_arabic 0.94 0.04 0.02 0.00    0
## relaxed_arabic     0.88 0.09 0.02 0.02    0
## visualise_arabic   0.91 0.04 0.04 0.01    0
```

L1 Arabic PCA

First, as previously outlined, it is important to assess whether particular items need to be excluded before continuing with the PCA. The same steps are followed as outlined with the L1 English data.

```
strategyArMatrix <- cor(strategy_ArQs)

#eyeball correlations
round(strategyArMatrix,2)
# "action", "context", "ignore_arabic" and "ignore_english" have no
correlations with other variables >.3

#check it's not an identity matrix - appears to be significant
cortest.bartlett(strategyArMatrix, n=109)

KMO(strategyArMatrix)
# KMO = .62 which is mediocre, mental_image, relaxed_english,
ignore_Arabic, and ignore_English are all <.5

#start by excluding these variables as they also had low correlations in
general

## rerun checks without excluded variables
strategy_ArQs2 <- strategy_ArQs %>%
  select(-c("ignore_arabic", "ignore_english", "mental_image",
"relaxed_english", "action", "context"))

strategyArMatrix2 <- cor(strategy_ArQs2)
corrplot(strategyArMatrix2,order="hclust", type="upper", tl.col="black",
tl.srt=45)

#eyeball correlations
round(strategyArMatrix2,2)
# "associations", "visualise_english" have have no correlations with
other variables >.3

#check it's not an identity matrix - appears to be significant
cortest.bartlett(strategyArMatrix2, n=109)

KMO(strategyArMatrix2)
# KMO = .69 which is better, all variable >.5

## rerun checks without excluded variables
strategy_ArQs2 <- strategy_ArQs2 %>%
  select(-c("associations", "visualise_english"))

strategyArMatrix2 <- cor(strategy_ArQs2)
corrplot(strategyArMatrix2,order="hclust", type="upper", tl.col="black",
tl.srt=45)
```

```
#eyeball correlations
round(strategyArMatrix2,2)
# all have at least 1 correlation >.3

#check it's not an identity matrix - appears to be significant
cortest.bartlett(strategyArMatrix2, n=109)

KMO(strategyArMatrix2)
# KMO = .7 which ok and all variables are over .5

det(strategyArMatrix2)
# determinant is greater than necessary value, so not problematic
```

After making the necessary changes, such as excluding the strategies outlined above, the preliminary analysis is satisfactory for the L1 Arabic data. The correlation matrix for the remaining items is provided in Figure XX.3.

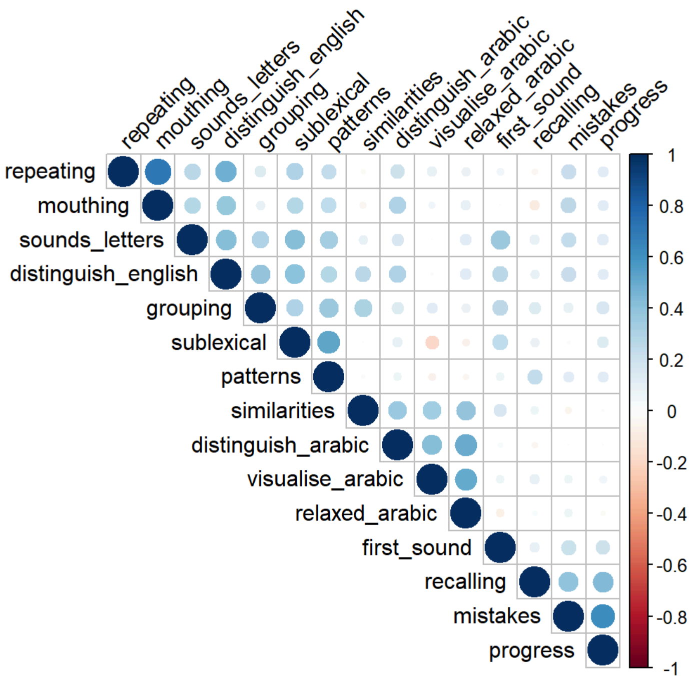


Figure XX.3: Correlation plot of all strategy items included in the L1 Arabic PCA analysis

As before, I use the principal() function and extract the same number of factors as there are variables, to inspect their eigenvalues and make further decisions about factor extraction.

```
#PCA model 1 with same number of factors as variables

pc1_ar <-principal(strategyArMatrix2, nfactors =
length(strategyArMatrix2[,1]), rotate = "none")
```

```
pc1_ar
# 5 factors have eigenvalues greater than 1 (Kaiser's criterion)

plot(pc1_ar$values, type = "b")
# this demonstrates evidence of a plateau after 6 factors...but
subjective

#additional analysis to consider how many factors to extract
results_Ar <- nScree(x=pc1_ar$value)
results_Ar
summary(results_Ar)
plotnScree(results_Ar)
```

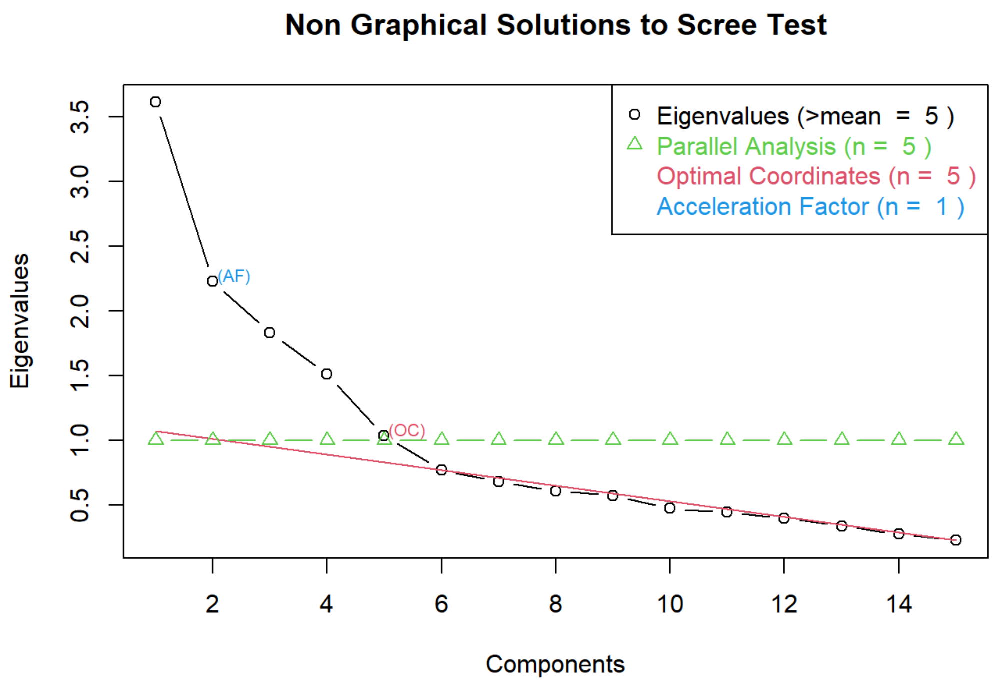


Figure XX.4: Scree plot including Parallel Analysis and Optimal Coordinates to assess number of factors to extract for the L1 Arabic data

Kaiser's criterion of including factors with eigenvalues over 1, as well as looking at the scree plot through the lens of parallel analysis and optimal coordinates, provides grounds to continue with PCA extracting 5 factors. This decision is additionally assessed by examining communalities of variables and residuals.

```
# continue with analysis extracting 7 factors
pc2_ar <-principal(strategyArMatrix2, nfactors = 5, rotate = "none")
pc2_ar
```

```
mean(pc2_ar$communality)

#commonality (h2 column) average is not greater than .7, which does not
meet Kaiser's criterion (=.68) - might be worth re-running analysis with
additional factors. Also, fit = .92 - above .9 as a min

#we can further check this by looking at residuals
residuals_ar2 <-factor.residuals(strategyArMatrix2, pc2_ar$loadings)
residuals_ar2 <-as.matrix(residuals_ar2[upper.tri(residuals_ar2)])
large.resid_ar2 <-abs(residuals_ar2)>0.05

sum(large.resid_ar2)
sum(large.resid_ar2)/nrow(large.resid_ar2)
#less than 50% of residuals greater than 0.05, so .49 is not ideal

sqrt(mean(residuals_ar2^2))
#quite high

hist(residuals_ar2)
#approximately normal
```

Re-run analysis with 6 factors extracted to check improved fit.

```
# continue with analysis extracting 6 factors
pc3_ar <-principal(strategyArMatrix2, nfactors = 6, rotate = "none")
pc3_ar

mean(pc3_ar$communality)

#commonality (h2 column) average is greater than .7, meeting Kaiser's
criterion (=.73) - fit = .93

#we can further check this by looking at residuals
residuals_ar3 <-factor.residuals(strategyArMatrix2, pc3_ar$loadings)
residuals_ar3 <-as.matrix(residuals_ar3[upper.tri(residuals_ar3)])
large.resid_ar3 <-abs(residuals_ar3)>0.05

sum(large.resid_ar3)
sum(large.resid_ar3)/nrow(residuals_ar3)
#less than 50% of residuals greater than 0.05, so this is ok - .36 is
better

sqrt(mean(residuals_ar3^2))
```

```
#quite high but better

hist(residuals_ar3)
#slight skew
```

Once the appropriate checks are satisfied, the PCA continues with the extraction of 6 factors. As before, to improve interpretation of factors and assuming that underlying dimensions are likely to be related to each other, an oblique rotation is applied to the data.

```
# oblique rotation

pc6_ar <- principal(strategy_ArQs2, nfactors = 6, rotate = "oblimin")
pc6_ar

print.psych(pc6_ar, cut = 0.3, sort = TRUE)
```

```
## Principal Components Analysis
## Call: principal(r = strategy_ArQs2, nfactors = 6, rotate = "oblimin")
## Standardized loadings (pattern matrix) based upon correlation matrix
##                     item   TC4   TC2   TC3   TC1   TC5   TC6   h2   u2 com
## repeating              9  0.89                                0.79 0.21 1.0
## mouthing              10  0.88                                0.81 0.19 1.1
## distinguish_english    5  0.52                          0.46 0.68 0.32 2.5
## relaxed_arabic        15        0.86                          0.73 0.27 1.0
## visualise_arabic      11        0.75                          0.65 0.35 1.2
## distinguish_arabic     4        0.72                          0.66 0.34 1.3
## progress              13              0.83                    0.72 0.28 1.0
## mistakes              12              0.81                    0.79 0.21 1.3
## recalling             14 -0.31        0.73  0.32              0.71 0.29 1.9
## patterns               8                    0.85              0.77 0.23 1.1
## sublexical             2                    0.70              0.71 0.29 1.5
## first_sound            3                          0.87        0.80 0.20 1.2
## sounds_letters         1                    0.40  0.68        0.73 0.27 1.9
## grouping               7                                0.76 0.71 0.29 1.2
## similarities           6        0.37                    0.69 0.73 0.27 1.7
##
##                        TC4  TC2  TC3  TC1  TC5  TC6
## SS loadings           2.18 2.10 2.00 1.76 1.49 1.46
## Proportion Var        0.15 0.14 0.13 0.12 0.10 0.10
## Cumulative Var        0.15 0.29 0.42 0.54 0.64 0.73
## Proportion Explained  0.20 0.19 0.18 0.16 0.14 0.13
## Cumulative Proportion 0.20 0.39 0.57 0.73 0.87 1.00
##
##  With component correlations of
##      TC4  TC2  TC3  TC1  TC5  TC6
## TC4 1.00  0.11 0.06  0.23 0.20 0.03
```

```
## TC2 0.11  1.00 0.02 -0.04 0.01 0.21
## TC3 0.06  0.02 1.00  0.11 0.14 0.08
## TC1 0.23 -0.04 0.11  1.00 0.16 0.16
## TC5 0.20  0.01 0.14  0.16 1.00 0.22
## TC6 0.03  0.21 0.08  0.16 0.22 1.00
##
## Mean item complexity =  1.4
## Test of the hypothesis that 6 components are sufficient.
##
## The root mean square of the residuals (RMSR) is  0.06
##  with the empirical chi square  94.41  with prob <  1.4e-08
##
## Fit based upon off diagonal values = 0.93
```

```
# should report the structure matrix as well as the pattern matrix above
(but a little hard to make sense of) - look for loadings of >.4 for both
matrices.

pc6_arStr<- pc6_ar$loadings %*% pc6_ar$Phi
pc6_arStr
```

```
##                               TC4         TC2         TC3         TC1
## sounds_letters       0.318275025  0.15551713  0.13518235  0.49177303
## sublexical           0.340463309 -0.10180323  0.01323417  0.77523557
## first_sound          0.027730199 -0.03942601  0.20506529  0.01659169
## distinguish_arabic   0.320416706  0.75460762 -0.08081845  0.11574219
## distinguish_english  0.592920767  0.11737708  0.10850849  0.32510096
## similarities        -0.073336244  0.50788791 -0.03865362 -0.06424413
## grouping             0.125387714  0.07200870  0.16069426  0.38436237
## patterns             0.221816629 -0.02008835  0.16818897  0.86629977
## repeating            0.885215371  0.11213145  0.11298419  0.22405088
## mouthing             0.889190020  0.16107842  0.09321807  0.24019215
## visualise_arabic     0.006771387  0.76804787  0.14398387 -0.19043552
## mistakes             0.295714595  0.04065068  0.82832043  0.01375827
## progress             0.126517899 -0.02496173  0.84348902  0.08637571
## recalling           -0.198519531  0.07658904  0.73272195  0.31632627
## relaxed_arabic       0.084594004  0.85383158  0.01058405 -0.01988546
##                               TC5          TC6
## sounds_letters       0.730985611  0.108760121
## sublexical           0.391844826  0.181620578
## first_sound          0.862813561  0.292488666
## distinguish_arabic   0.098487428  0.230462489
## distinguish_english  0.445922715  0.527920822
## similarities         0.140263028  0.745469299
## grouping             0.254483190  0.793552690
## patterns             0.076370186  0.226680680
## repeating            0.151006116  0.074815780
```

```
## mouthing              0.114518685 -0.046825961
## visualise_arabic     -0.001844755  0.219329421
## mistakes              0.270830482 -0.002700903
## progress              0.179854823  0.098169327
## recalling            -0.008257398  0.162313826
## relaxed_arabic       -0.002418980  0.150007683
```

TC1 and TC4 are marginally correlated (.23), which is logical in relation to the connection found between using OI as a basis for production. Next, it is necessary to conduct reliability analyses on subscales which have at least three items which have a loading >.4 and do not load highly across multiple other factors. There are four subscales that meet this criteria:

TC4 = Learning by producing
TC2 = Using Arabic OI
TC3 = Evaluating learning
TC1 = Analysing sounds and letters

```
producing <- strategy_ArQs2 [, c("repeating", "mouthing",
"distinguish_english")]

usingArabicOI <- strategy_ArQs2 [, c("relaxed_arabic",
"visualise_arabic", "distinguish_arabic")]

evaluating<- strategy_ArQs2[, c("recalling", "progress", "mistakes")]

analysing <- strategy_ArQs2 [, c("patterns", "sublexical",
"sounds_letters")]
```

Calculating the alpha is straightforward for all, as they are all positively coded items.

```
alpha(producing)
#overall reliability is good (alpha =.77), reliability would increase
without "distinguish_english" but r.drop are all above .3.
```

```
##
## Reliability analysis
## Call: alpha(x = producing)
##
##   raw_alpha std.alpha G6(smc) average_r S/N   ase mean  sd median_r
##       0.77      0.76    0.72      0.52 3.2 0.039  3.2 1.1     0.48
##
##     95% confidence boundaries
```

```
##          lower alpha upper
## Feldt     0.68  0.77  0.83
## Duhachek  0.69  0.77  0.84
##
##  Reliability if an item is dropped:
##                     raw_alpha std.alpha G6(smc) average_r S/N alpha se var.r
## repeating                0.55      0.55    0.38      0.38 1.2    0.087    NA
## mouthing                 0.65      0.65    0.48      0.48 1.9    0.067    NA
## distinguish_english      0.82      0.82    0.70      0.70 4.7    0.034    NA
##                     med.r
## repeating            0.38
## mouthing             0.48
## distinguish_english  0.70
##
##  Item statistics
##                       n raw.r std.r r.cor r.drop mean  sd
## repeating           109  0.89  0.88  0.82   0.72  3.4 1.4
## mouthing            109  0.84  0.84  0.75   0.63  3.3 1.3
## distinguish_english 109  0.74  0.75  0.52   0.47  2.8 1.3
##
## Non missing response frequency for each item
##                        1    2    3    4    5 miss
## repeating           0.16 0.11 0.22 0.25 0.27    0
## mouthing            0.15 0.09 0.29 0.22 0.25    0
## distinguish_english 0.18 0.26 0.26 0.19 0.11    0
```

```
alpha(usingArabicOI)
#overall reliability is good (alpha =.73), none of the items would
increase reliability if they were deleted and the r.drop values are
above .3.
```

```
##
## Reliability analysis
## Call: alpha(x = usingArabicOI)
##
##   raw_alpha std.alpha G6(smc) average_r S/N   ase mean sd median_r
##       0.73      0.73    0.65      0.48 2.7 0.045  2.4  1      0.5
##
##     95% confidence boundaries
##          lower alpha upper
## Feldt     0.63  0.73  0.81
## Duhachek  0.64  0.73  0.82
##
##  Reliability if an item is dropped:
##                    raw_alpha std.alpha G6(smc) average_r S/N alpha se var.r
## relaxed_arabic          0.59      0.59    0.42      0.42 1.5    0.078    NA
## visualise_arabic        0.66      0.67    0.50      0.50 2.0    0.064    NA
```

```
## distinguish_arabic      0.67      0.67    0.50      0.50 2.0    0.063    NA
##                    med.r
## relaxed_arabic      0.42
## visualise_arabic    0.50
## distinguish_arabic  0.50
##
##  Item statistics
##                      n raw.r std.r r.cor r.drop mean  sd
## relaxed_arabic     109  0.83  0.83  0.70   0.59  2.7 1.3
## visualise_arabic   109  0.81  0.80  0.63   0.54  2.3 1.3
## distinguish_arabic 109  0.78  0.79  0.62   0.53  2.1 1.2
##
## Non missing response frequency for each item
##                       1    2    3    4    5 miss
## relaxed_arabic     0.18 0.29 0.27 0.11 0.15    0
## visualise_arabic   0.37 0.23 0.19 0.12 0.09    0
## distinguish_arabic 0.39 0.31 0.17 0.06 0.06    0
```

```
alpha(evaluating)
#overall reliability is good (alpha = .73), reliability would increase
without "recalling" but the r.drop values are all above .3
```

```
##
## Reliability analysis
## Call: alpha(x = evaluating)
##
##   raw_alpha std.alpha G6(smc) average_r S/N   ase mean   sd median_r
##       0.73      0.74    0.67      0.48 2.8 0.045  3.4 0.85     0.44
##
##     95% confidence boundaries
##          lower alpha upper
## Feldt     0.63  0.73  0.81
## Duhachek  0.65  0.73  0.82
##
##  Reliability if an item is dropped:
##           raw_alpha std.alpha G6(smc) average_r S/N alpha se var.r med.r
## recalling      0.76      0.76    0.61      0.61 3.2    0.046    NA  0.61
## progress       0.56      0.57    0.40      0.40 1.3    0.083    NA  0.40
## mistakes       0.61      0.61    0.44      0.44 1.6    0.075    NA  0.44
##
##  Item statistics
##             n raw.r std.r r.cor r.drop mean   sd
## recalling 109  0.77  0.76  0.53   0.47  3.3 1.09
## progress  109  0.85  0.84  0.74   0.62  3.3 1.09
## mistakes  109  0.81  0.83  0.71   0.59  3.6 0.96
##
## Non missing response frequency for each item
##              1    2    3    4    5 miss
```

```
## recalling 0.05 0.18 0.38 0.23 0.17    0
## progress  0.06 0.17 0.33 0.30 0.15    0
## mistakes  0.02 0.10 0.30 0.39 0.18    0
```

```
alpha(analysing)
#overall reliability is ok but not great (alpha = .68). There is some
evidence that reliability would be slightly improved by dropping
"sounds_letters" from this subscale but the r.drop values are all above
.3.
```

```
##
## Reliability analysis
## Call: alpha(x = analysing)
##
##   raw_alpha std.alpha G6(smc) average_r S/N   ase mean   sd median_r
##       0.68      0.68     0.6      0.42 2.2 0.053  2.8 0.98     0.41
##
##     95% confidence boundaries
##          lower alpha upper
## Feldt     0.57  0.68  0.77
## Duhachek  0.58  0.68  0.79
##
##  Reliability if an item is dropped:
##                raw_alpha std.alpha G6(smc) average_r  S/N alpha se var.r
med.r
## patterns            0.58      0.58    0.41      0.41 1.39    0.080    NA
0.41
## sublexical          0.49      0.49    0.33      0.33 0.98    0.097    NA
0.33
## sounds_letters      0.69      0.69    0.52      0.52 2.18    0.060    NA
0.52
##
##  Item statistics
##                  n raw.r std.r r.cor r.drop mean  sd
## patterns       109  0.79  0.79  0.62   0.51  2.8 1.3
## sublexical     109  0.82  0.82  0.69   0.57  2.2 1.2
## sounds_letters 109  0.74  0.74  0.51   0.42  3.4 1.3
##
## Non missing response frequency for each item
##                   1    2    3    4    5 miss
## patterns       0.20 0.19 0.35 0.12 0.14    0
## sublexical     0.38 0.28 0.18 0.08 0.07    0
## sounds_letters 0.11 0.13 0.26 0.30 0.20    0
```

Factor scores for participants are calculated and added to a dataframe with participant IDs. Next, these scores are compiled with other demographic information and performance data from the matching task for further analysis.

```
pc7_ar <-principal(strategy_ArQs2, nfactors = 6, rotate = "oblimin",
scores = TRUE)

pc7_ar$scores
```

```
# add scores to data frame with participant IDs and correct scores.

strategy_scoresAr <-cbind(strategy_ID_ar,pc7_ar$scores)

# rename 4 key subscales

names(strategy_scoresAr)[28]<- "producing"
names(strategy_scoresAr)[29]<- "usingArabicOI"
names(strategy_scoresAr)[30]<- "evaluating"
names(strategy_scoresAr)[31]<- "analysing"

#select the key variables of interest

strategy_scoresAr2 <- select(strategy_scoresAr, ID, producing,
usingArabicOI, evaluating, analysing)

#select the key demographic variables, as well as accuracy and RT data
from the audiovisual matching task, just for the L1 Arabic group.

strategy_ID_short2 <- select(strategy_rename, ID, L1, age, gender,
education, nationality, readingprofscore, prof_test_score, level,
prop_English_lit_hours, prop_arabic_lit_hours, meanRT, meanD.prime,
correct, prop_distraction)

strategy_ID_shortAr <- filter(strategy_ID_short2, L1 == "Arabic")

#join to create a combined df with key information

strategy_scoresAr2 <- inner_join(strategy_ID_shortAr,
strategy_scoresAr2, by = "ID")
```

Appendix XXI: LLS correlational and regression Rmd output

Correlations between the strategy scores, accuracy data and individual differences data are calculated using Spearman's nonparametric rank-order correlation. Individual difference (ID) variables that have been significant in previous analysis are included here, namely: age, proficiency and exposure to English literacy. *d′* is used as the measure of accuracy, rather than raw correct scores, as a more insightful measure of sensitivity and precision.

```
arCorr <- strategy_scoresAr2 %>%
  select(age, prop_distraction, prof_test_score, prop_English_lit_hours,
meanD.prime, producing, usingArabicOI, evaluating, analysing)

ArStrat = cor(arCorr, method = "s", use = "complete.obs")

# adding confidence intervals and p values
testRes = cor.mtest(arCorr, conf.level = 0.95)

#leave blank non-significant coefficients and add significant
correlation coefficient
corrplot(ArStrat, p.mat=testRes$p,
         method = 'circle', tl.col = "black", type = 'lower',
         insig = 'blank', addCoef.col = 'black',
         number.cex = 0.8, order = 'AOE', diag = FALSE, )
```

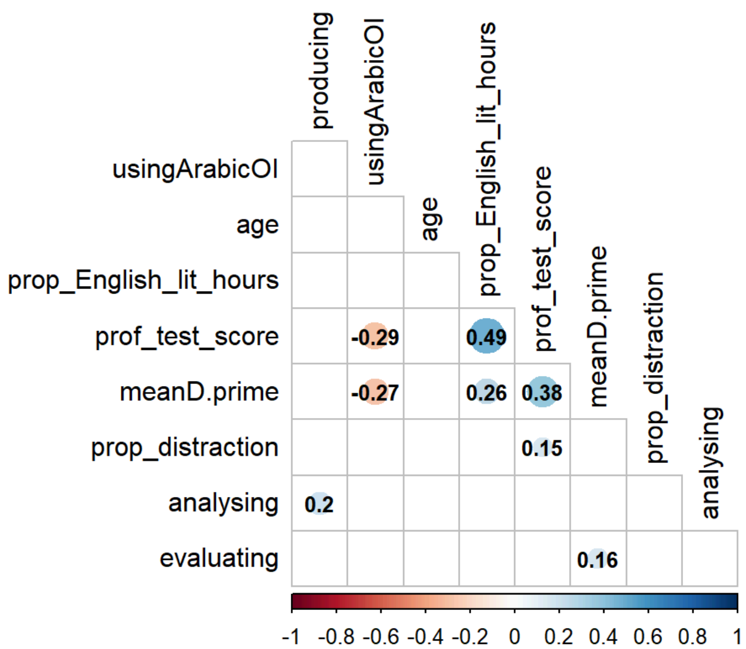


Figure XXI.1: Correlations between L1 Arabic LLS, task performance and IDs

Relationships are further explored with regression analyses, to investigate whether strategy use predicts increased accuracy in the matching task. The first approach uses multiple regression analysis to explore the relationship between proficiency, LLS and accuracy.

```
model1 <- lm(meanD.prime ~ prof_test_score, data = strategy_scoresAr2)
summary(model1)
```

```
##
## Call:
## lm(formula = meanD.prime ~ prof_test_score, data = strategy_scoresAr2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.24181 -0.43977 -0.00279  0.33421  1.43164
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.04367    0.25769   0.169    0.866
## prof_test_score  0.10878    0.02612   4.164 6.35e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5897 on 107 degrees of freedom
## Multiple R-squared:  0.1394, Adjusted R-squared:  0.1314
## F-statistic: 17.34 on 1 and 107 DF,  p-value: 6.355e-05
```

```
model2 <- lm(meanD.prime ~ prof_test_score + age + prop_distraction,
data = strategy_scoresAr2)
summary(model2)
```

```
##
## Call:
## lm(formula = meanD.prime ~ prof_test_score + age + prop_distraction,
##     data = strategy_scoresAr2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.26315 -0.36659 -0.04166  0.37548  1.39997
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)       0.172582   0.363084   0.475    0.636
## prof_test_score   0.114570   0.026732   4.286 4.05e-05 ***
## age              -0.003786   0.007447  -0.508    0.612
## prop_distraction -0.283559   0.214196  -1.324    0.188
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5899 on 105 degrees of freedom
## Multiple R-squared:  0.155,  Adjusted R-squared:  0.1309
## F-statistic: 6.421 on 3 and 105 DF,  p-value: 0.000491
```

```
anova(model1, model2)

# exposure to English literacy and proportion of distraction do not
contribute much to the model - not significant, don't reduce standard
error and only marginally improve R^2
```

```
## Analysis of Variance Table
##
## Model 1: meanD.prime ~ prof_test_score
## Model 2: meanD.prime ~ prof_test_score + age + prop_distraction
##   Res.Df    RSS Df Sum of Sq      F Pr(>F)
## 1    107 37.214
## 2    105 36.540  2   0.67362 0.9678 0.3833
```

```
# run again with strategy scores but keeping proficiency

model3 <- lm(meanD.prime ~ prof_test_score + producing + usingArabicOI +
evaluating + analysing, data = strategy_scoresAr2)
summary(model3)
```

```
##
## Call:
## lm(formula = meanD.prime ~ prof_test_score + producing + usingArabicOI +
##     evaluating + analysing, data = strategy_scoresAr2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.26290 -0.37652 -0.02673  0.38109  1.36625
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.18495    0.26865   0.688 0.492728
## prof_test_score  0.09410    0.02732   3.445 0.000828 ***
## producing       -0.02960    0.05841  -0.507 0.613357
## usingArabicOI   -0.09433    0.05818  -1.621 0.108012
## evaluating       0.13030    0.05597   2.328 0.021866 *
## analysing        0.02191    0.05814   0.377 0.707004
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5777 on 103 degrees of freedom
## Multiple R-squared:  0.2052, Adjusted R-squared:  0.1666
## F-statistic: 5.318 on 5 and 103 DF,  p-value: 0.0002183
```

```
anova(model1, model3)
# approaching significance and improves model fit
```

```
## Analysis of Variance Table
##
## Model 1: meanD.prime ~ prof_test_score
## Model 2: meanD.prime ~ prof_test_score + producing + usingArabicOI +
evaluating +
##     analysing
##   Res.Df    RSS Df Sum of Sq      F  Pr(>F)
## 1    107 37.214
## 2    103 34.371  4    2.8427 2.1297 0.08239 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

A number of checks are then applied to make sure necessary assumptions were met.

```
strategy_scoresAr2$residuals<-resid(model3)
strategy_scoresAr2$standardised.residuals<-rstandard(model3)
strategy_scoresAr2$cooks.distance<-cooks.distance(model3)
strategy_scoresAr2$dffit<-dffits(model3)
strategy_scoresAr2$leverage<-hatvalues(model3)
strategy_scoresAr2$covariance<-covratio(model3)

write.table(strategy_scoresAr2, "strategy_accuragy_diagnostics", sep = "\t",
row.names = FALSE)

strategy_scoresAr2$standardised.residuals >
2|strategy_scoresAr2$standardised.residuals < -2

strategy_scoresAr2$large.residual <- strategy_scoresAr2$standardised.residuals >
2|strategy_scoresAr2$standardised.residuals < -2

sum(strategy_scoresAr2$large.residual)

# closer look at large residuals reveals that 5 of the residuals lie outside of
+_ 2.5 - but looks ok
# cooks distance looks fine (none greater than 1)
# seems ok all taken together

strategy_scoresAr2[strategy_scoresAr2$large.residual,
```

```r
c("standardised.residuals", "cooks.distance", "leverage", "covariance")]
```

```
##    standardised.residuals cooks.distance   leverage covariance
## 20               2.405028     0.03325566 0.03334629  0.7754437
## 29              -2.103708     0.01693927 0.02244989  0.8333840
## 68              -2.225138     0.05086253 0.05805766  0.8375819
## 81              -2.219679     0.02533887 0.02993366  0.8145073
## 84               2.387122     0.01776706 0.01836403  0.7676613
```

```r
## assessing independence
dwt(model3)
# good - DWT statistic is close to 2
```

```
##  lag Autocorrelation D-W Statistic p-value
##    1       -0.106716      2.192209   0.326
##  Alternative hypothesis: rho != 0
```

```r
## assessing multicolinearity
vif(model3)
# looks good (none greater than 10)
```

```
## prof_test_score       producing    usingArabicOI       evaluating        analysing
##        1.139475        1.104144         1.095627         1.013841         1.093947
```

```r
1/vif(model3)
# looks good - no tolerances near 2
```

```
## prof_test_score       producing    usingArabicOI       evaluating        analysing
##       0.8775971       0.9056790        0.9127197        0.9863479        0.9141214
```

```r
mean(vif(model3))
# close to 1 - is good
```

```
## [1] 1.089407
```

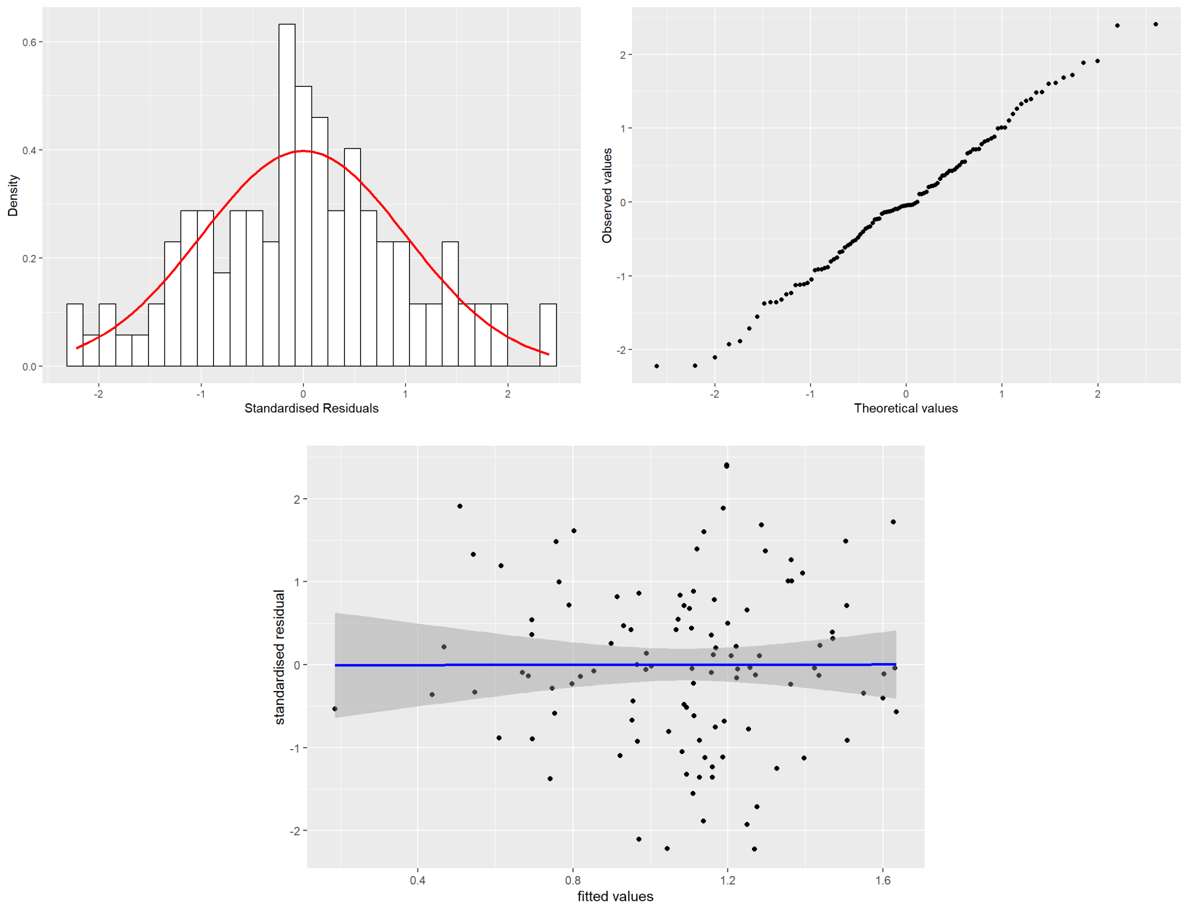


Figure XXI.2: Multiple regression residuals

Next, linear regressions are used to investigate whether proficiency and exposure to English literacy predict strategy useage. First, learning by producing new words:

```
# producing
model7 <-lm(producing ~ prof_test_score, data = strategy_scoresAr2)
summary(model7)
```

```
##
## Call:
## lm(formula = producing ~ prof_test_score, data = strategy_scoresAr2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.94767 -0.59207 -0.03353  0.77691  1.88455
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.73382    0.43293   1.695   0.0930 .
## prof_test_score -0.07625    0.04389  -1.737   0.0852 .
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9908 on 107 degrees of freedom
## Multiple R-squared:  0.02743,    Adjusted R-squared:  0.01834
## F-statistic: 3.018 on 1 and 107 DF,  p-value: 0.08522
```

```
model7b <-lm(producing ~ prof_test_score + prop_English_lit_hours, data =
strategy_scoresAr2)

summary(model7b)
```

```
## Call:
## lm(formula = producing ~ prof_test_score + prop_English_lit_hours,
##     data = strategy_scoresAr2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.00840 -0.51622 -0.08808  0.67830  1.68051
##
## Coefficients:
##                         Estimate Std. Error t value Pr(>|t|)
## (Intercept)              0.79586    0.43589   1.826   0.0707 .
## prof_test_score         -0.10267    0.04973  -2.065   0.0414 *
## prop_English_lit_hours  1.83599    1.63163   1.125   0.2630
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9896 on 106 degrees of freedom
## Multiple R-squared:  0.03891,    Adjusted R-squared:  0.02078
## F-statistic: 2.146 on 2 and 106 DF,  p-value: 0.122
```

```
anova(model7, model7b)
#doesn't improve fit
```

```
## Analysis of Variance Table
##
## Model 1: producing ~ prof_test_score
## Model 2: producing ~ prof_test_score + prop_English_lit_hours
##   Res.Df    RSS Df Sum of Sq      F Pr(>F)
## 1    107 105.04
## 2    106 103.80  1    1.2399 1.2662  0.263
```

Second, learning by using Arabic written input:

```
# using Arabic OI
```

```
model8 <-lm(usingArabicOI ~ prof_test_score, data = strategy_scoresAr2)
summary(model8)
```

```
##
## Call:
## lm(formula = usingArabicOI ~ prof_test_score, data = strategy_scoresAr2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.07092 -0.68181 -0.01708  0.56722  2.35089
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.27810    0.42033   3.041  0.00297 **
## prof_test_score -0.13281    0.04261  -3.116  0.00235 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.962 on 107 degrees of freedom
## Multiple R-squared:  0.08322,    Adjusted R-squared:  0.07465
## F-statistic: 9.713 on 1 and 107 DF,  p-value: 0.00235
```

```
model8b <-lm(usingArabicOI ~ prof_test_score + prop_English_lit_hours, data =
strategy_scoresAr2)
summary(model8b)
```

```
##
## Call:
## lm(formula = usingArabicOI ~ prof_test_score + prop_English_lit_hours,
##     data = strategy_scoresAr2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.04555 -0.69791 -0.00756  0.54991  2.22576
##
## Coefficients:
##                        Estimate Std. Error t value Pr(>|t|)
## (Intercept)             1.31614    0.42472   3.099  0.00249 **
## prof_test_score        -0.14900    0.04845  -3.075  0.00268 **
## prop_English_lit_hours  1.12590    1.58982   0.708  0.48038
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9642 on 106 degrees of freedom
## Multiple R-squared:  0.08754,    Adjusted R-squared:  0.07032
## F-statistic: 5.084 on 2 and 106 DF,  p-value: 0.007789
```

```
anova(model8, model8b)
#doesn't significantly improve fit
```

```
## Analysis of Variance Table
##
## Model 1: usingArabicOI ~ prof_test_score
## Model 2: usingArabicOI ~ prof_test_score + prop_English_lit_hours
##   Res.Df    RSS Df Sum of Sq      F Pr(>F)
## 1    107 99.012
## 2    106 98.546  1   0.46627 0.5015 0.4804
```

Third, evaluating learning:

```
# evaluating
model9 <-lm(evaluating ~ prof_test_score, data = strategy_scoresAr2)
summary(model9)
```

```
##
## Call:
## lm(formula = evaluating ~ prof_test_score, data = strategy_scoresAr2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3.10336 -0.64071  0.00522  0.68286  2.08046
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.095622   0.438890   0.218    0.828
## prof_test_score -0.009936   0.044495  -0.223    0.824
##
## Residual standard error: 1.004 on 107 degrees of freedom
## Multiple R-squared:  0.0004658,  Adjusted R-squared:  -0.008876
## F-statistic: 0.04986 on 1 and 107 DF,  p-value: 0.8237
```

```
model9b <-lm(evaluating ~ prof_test_score + prop_English_lit_hours, data =
strategy_scoresAr2)
summary(model9b)
```

```
##
## Call:
## lm(formula = evaluating ~ prof_test_score + prop_English_lit_hours,
##     data = strategy_scoresAr2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3.10763 -0.62732 -0.00706  0.69836  2.08465
```

```
##
## Coefficients:
##                         Estimate Std. Error t value Pr(>|t|)
## (Intercept)             0.082887   0.444418   0.187    0.852
## prof_test_score        -0.004513   0.050699  -0.089    0.929
## prop_English_lit_hours -0.376906   1.663540  -0.227    0.821
##
## Residual standard error: 1.009 on 106 degrees of freedom
## Multiple R-squared:  0.0009496,  Adjusted R-squared:  -0.0179
## F-statistic: 0.05038 on 2 and 106 DF,  p-value: 0.9509
```

```
anova(model9, model9b)
#doesn't significantly improve fit
```

```
## Analysis of Variance Table
##
## Model 1: evaluating ~ prof_test_score
## Model 2: evaluating ~ prof_test_score + prop_English_lit_hours
##   Res.Df    RSS Df Sum of Sq      F Pr(>F)
## 1    107 107.95
## 2    106 107.90  1  0.052252 0.0513 0.8212
```

Fourth, learning by analysing sounds and letters:

```
# analysing
model10 <-lm(analysing ~ prof_test_score, data = strategy_scoresAr2)
summary(model10)
```

```
##
## Call:
## lm(formula = analysing ~ prof_test_score, data = strategy_scoresAr2)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.6875 -0.7000 -0.0951  0.5715  2.5986
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)     -0.52220    0.43593  -1.198    0.234
## prof_test_score  0.05426    0.04420   1.228    0.222
##
## Residual standard error: 0.9977 on 107 degrees of freedom
## Multiple R-squared:  0.01389,    Adjusted R-squared:  0.004676
## F-statistic: 1.507 on 1 and 107 DF,  p-value: 0.2222
```

```
model10b <-lm(analysing ~ prof_test_score + prop_English_lit_hours, data =
strategy_scoresAr2)
summary(model10b)
```

```
##
## Call:
## lm(formula = analysing ~ prof_test_score + prop_English_lit_hours,
##     data = strategy_scoresAr2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.72482 -0.75401 -0.04805  0.54985  2.57825
##
## Coefficients:
##                        Estimate Std. Error t value Pr(>|t|)
## (Intercept)            -0.56008    0.44057  -1.271    0.206
## prof_test_score         0.07039    0.05026   1.400    0.164
## prop_English_lit_hours -1.12097    1.64914  -0.680    0.498
##
## Residual standard error: 1 on 106 degrees of freedom
## Multiple R-squared:  0.01817,    Adjusted R-squared:  -0.0003536
## F-statistic: 0.9809 on 2 and 106 DF,  p-value: 0.3783
```

```
anova(model10, model10b)
```

```
## Analysis of Variance Table
##
## Model 1: analysing ~ prof_test_score
## Model 2: analysing ~ prof_test_score + prop_English_lit_hours
##   Res.Df    RSS Df Sum of Sq     F Pr(>F)
## 1    107 106.50
## 2    106 106.04  1    0.4622 0.462 0.4982
```

Appendix XXII: Cluster analysis Rmd output

This document outlines the procedure for conducting a cluster analysis to explore how L1 Arabic participants can be grouped in relation to their proficiency level (cluster package). This analysis follows the steps laid out by Crowther et al. (2021).

Variable to cluster: L1 Arabic participants (aggregated data per person)
Predictor variables: Proficiency test score, self-reported level, self-reported skills ability (mean across speaking, listening, reading and writing), estimated daily exposure to English.

The data frame, including strategy scores, is corrected for data type and releveled, if necessary. Self-reported level is recoded as numeric (none = 1, beginner = 2, intermediate = 3, advanced = 4, nearnative = 5). Predictor variables differ in length, so need to be scaled.

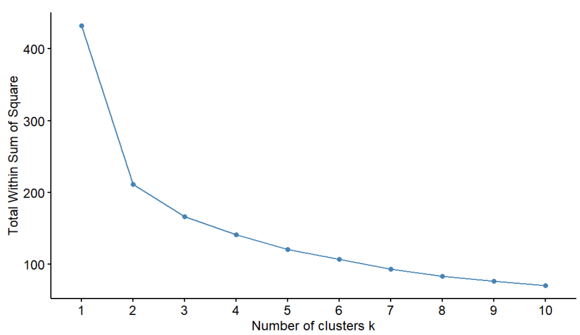
```
data1 <- data_raw %>% select (prof_test_score, mean_prof, level.n,
prop_English_hours) %>% scale()
```

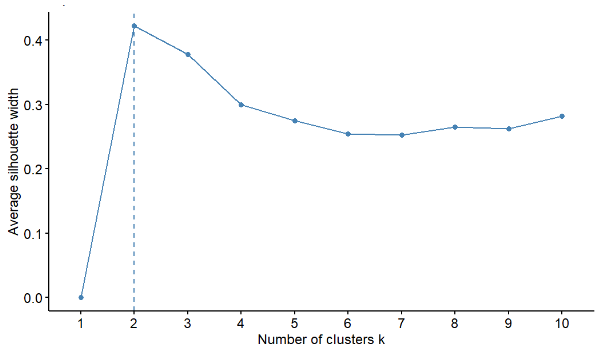
The first step involves determining the number of clusters. A few different indices are used to guide this decision, namely the Elbow method, Silhouette method and Gap statistic.

```
# Elbow method
fviz_nbclust(data1, hcut, method = "wss")

# Silhouette method
fviz_nbclust(data1, hcut, method = "silhouette")

# Gap statistic
fviz_nbclust(data1, hcut, method = "gap_stat", nboot = 500)
# note. nboot = 500 is recommended by (Kassambara, 2017)
```





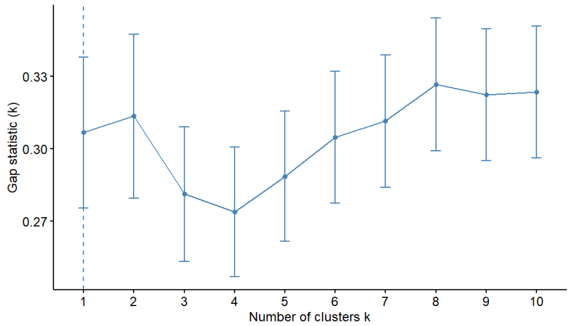


Figure XXII.1 Visualisations of optimal number of clusters using the (1) elbow method, (2) silhouette method, and (3) gap statistic, respectively.

These indices suggest that a 2-cluster solution may be optimal. To further investigate the validity of a two-cluster solution, a scatter plot for each cluster solution is provided, alongside statistical tests for group difference.

```
# Based on what the indices suggest, compare possible cluster solutions
# Here, we compare a 2-cluster, a 3-cluster and a 4-cluster solution

# Run and save a 2-cluster, 3-cluster and a 4-cluster solution
hca_cl2 <- eclust(data1, "hclust", k = 2, graph = FALSE)
hca_cl3 <- eclust(data1, "hclust", k = 3, graph = FALSE)

# Visualize
# 2-cluster solution
fviz_silhouette(hca_cl2) + theme_minimal()
plot2 <- fviz_cluster(hca_cl2) + theme_minimal()
fviz_dend(hca_cl2)

# 3-cluster solution
fviz_silhouette(hca_cl3) + theme_minimal()
plot3 <- fviz_cluster(hca_cl3) + theme_minimal()
fviz_dend(hca_cl3)
```

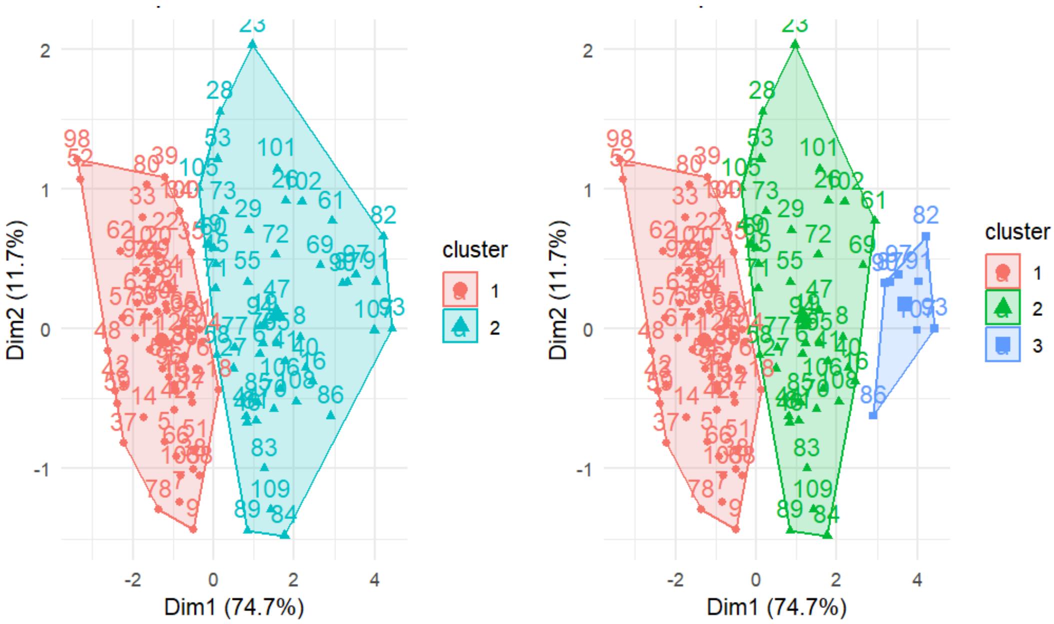


Figure XXII.2 Scatterplot of two and three-cluster solutions

```
# Decision: Final cluster solution as a 2-cluster
# 1) Analysis with 2-cluster solution

data_raw$cluster <- hca_cl2$cluster

data_raw$cluster <- as.factor(data_raw$cluster)

df_long <- gather(data_raw, "measure", "score", c(9, 10, 12, 23)) %>%
mutate(cluster = as.factor(cluster))

# Visually inspect the data
ggplot(df_long, aes(x = measure, y = score, fill = cluster)) +
 geom_boxplot(width = 0.5, alpha = .7, outlier.shape = 4) +
 stat_summary(aes(group = cluster), geom = "errorbar", width = .1, size
= 1,
 fun.data = "mean_cl_normal", color = "firebrick", position =
position_dodge(.5),
 show.legend = FALSE) +
 stat_summary(aes(group = cluster), geom = "point", fun.y = "mean",
color = "black", size = 2,
 position = position_dodge(.5),
 show.legend = FALSE) +
 #scale_y_continuous(limits = c(0,4500)) +
 theme_minimal(base_size = 14) #+ theme(legend.position = "none")
```

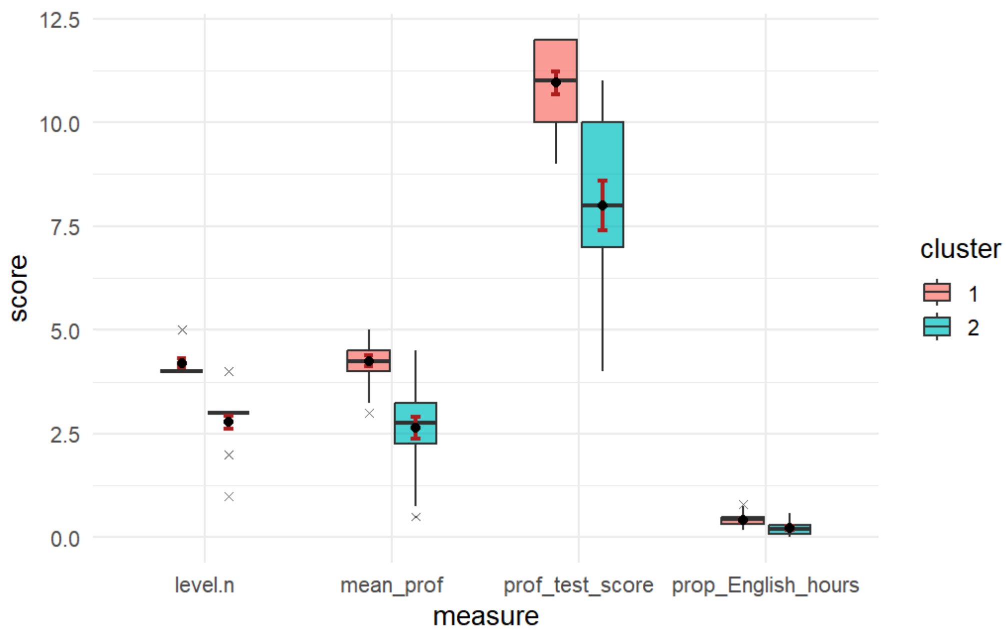


Figure XXII.3 Boxplots for the predictor variables by cluster

Cluster 1 appears to be higher proficiency and cluster 2 appears to be lower proficiency.

```
# 2) Group summaries for each cluster
data_raw%>% group_by(cluster) %>%
 summarize(profM = mean(prof_test_score), profSD = sd(prof_test_score),
           accM = mean(correct), accSD = sd(correct),
           dpriM = mean(meanD.prime), dpriSD = sd(meanD.prime),
           RTM = mean(meanRT), RTSD = sd(meanRT),
           prodM = mean(producing), prodSD = sd(producing),
           arOIM = mean(usingArabicOI), arOISD = sd(usingArabicOI),
           evalM = mean(evaluating), evalSD = sd(evaluating),
           analyseM = mean(analysing), analyseSD = sd(analysing)
           )

# See cluster composition
data_raw %>% group_by(cluster) %>% count(education)
```

Table XXII.1: Summary of variables by cluster, with either mean and standard deviation or number of participants

	Cluster 1	**Cluster 2**
Proficiency (*M*(*SD*))	10.95 (1.1)	8.00 (2.1)
Performance (*M*(*SD*))		
Accuracy	33.51 (5.9))	30.86 (6.6)

d′	1.22 (0.6)	0.93 (0.7)
RT	2093.36 (847.9)	2286.18 (877.2)
Strategies (*M*(*SD*))		
Producing	-0.16 (1.1)	0.20 (0.9)
Using Arabic OI	-0.15 (1.1)	0.18 (0.8)
Evaluating	0.02 (1.0)	-0.02 (1.0)
Analysing	0.13 (1.1)	-0.16 (0.9)
Education (*n*)		
Secondary	3	8
Professional qualification	2	0
Bachelors	14	21
Masters	13	29
Doctorate	12	7

Table XXIII.1 shows that cluster 1 has higher proficiency, higher accuracy in the lexical encoding task and faster reaction times than cluster 2. Additionally, cluster 1 uses less producing and Arabic written input strategies than cluster 2, meanwhile they use more analysing strategies. A mix of educational backgrounds are found across the two clusters.

Non-parametric Mann-Whitney U tests reveal that the means for the two clusters significantly differed across the L2 English proficiency and exposure measures. Also accuracy (raw scores and *d′*) and reliance on Arabic OI as a learning strategy significantly differed between the two clusters as well.

```
# 3) Cluster evaluation
# Mann-whitney test for cluster mean comparision
wilcox.test(prof_test_score ~ cluster, data = data_raw)
wilcox.test(mean_prof ~ cluster, data = data_raw)
wilcox.test(level.n ~ cluster, data = data_raw)
wilcox.test(prop_English_hours ~ cluster, data = data_raw)


# cluster evaluation of other variables
# Kruskal-Wallis test for cluster mean comparision
wilcox.test(correct ~ cluster, data = data_raw)
wilcox.test(meanD.prime ~ cluster, data = data_raw)
wilcox.test(meanRT ~ cluster, data = data_raw)
wilcox.test(producing ~ cluster, data = data_raw)
wilcox.test(usingArabicOI ~ cluster, data = data_raw)
wilcox.test(evaluating ~ cluster, data = data_raw)
```

```r
wilcox.test(analysing ~ cluster, data = data_raw)
```

```
Wilcoxon rank sum test with continuity correction
data:  prof_test_score by cluster
W = 2636.5, p-value = 5.28e-13
alternative hypothesis: true location shift is not equal to 0

data:  mean_prof by cluster
W = 2790, p-value = 7.189e-16
alternative hypothesis: true location shift is not equal to 0

data:  level.n by cluster
W = 2892, p-value < 2.2e-16
alternative hypothesis: true location shift is not equal to 0

data:  prop_English_hours by cluster
W = 2447, p-value = 2.683e-09
alternative hypothesis: true location shift is not equal to 0

data:  correct by cluster
W = 1864, p-value = 0.01634
alternative hypothesis: true location shift is not equal to 0

data:  meanD.prime by cluster
W = 1874, p-value = 0.01397
alternative hypothesis: true location shift is not equal to 0

data:  meanRT by cluster
W = 1230, p-value = 0.1446
alternative hypothesis: true location shift is not equal to 0

data:  producing by cluster
W = 1196, p-value = 0.09571
alternative hypothesis: true location shift is not equal to 0

data:  usingArabicOI by cluster
W = 1141, p-value = 0.04539
alternative hypothesis: true location shift is not equal to 0

data:  evaluating by cluster
W = 1475, p-value = 0.9781
alternative hypothesis: true location shift is not equal to 0

data:  analysing by cluster
W = 1704, p-value = 0.1549
alternative hypothesis: true location shift is not equal to 0
```

Variables of interest from chapter 5 are then combined with the data file with assigned clusters, to see how OI influence perceptions and beliefs relate to proficiency

```
debrief_ar <- debrief_data%>%
  filter(L1 == "Arabic")

OI_influence <- select(debrief_ar2, ID, see_OI, see_arabic, see_english,
no_spell, spelling.preference, OI.importance)

#inner_join rather than merge used as different length of datasets,
based on the exluded non-responses to the strategy inventory
prof_cluster <- inner_join(data_raw, OI_influence,  by = "ID")
```

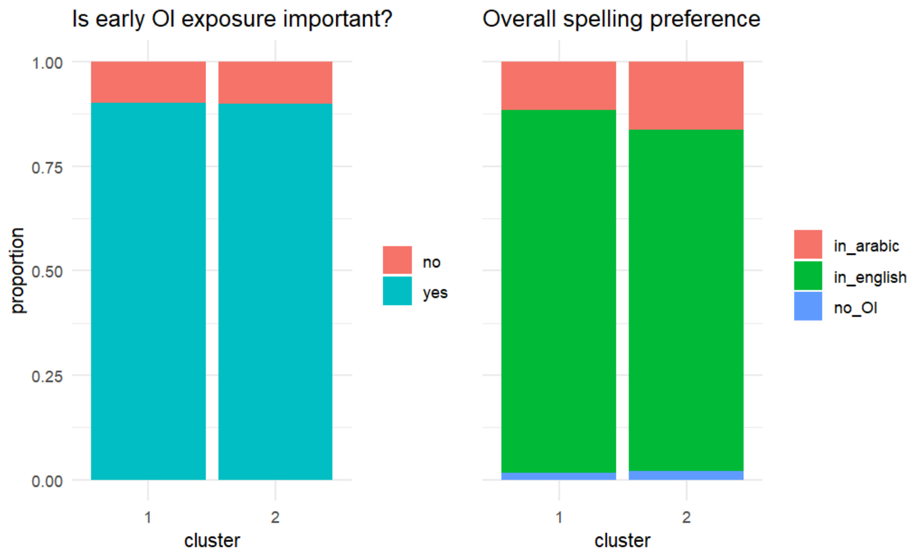


Figure XXII.4 Bar plots of participant perceptions by cluster

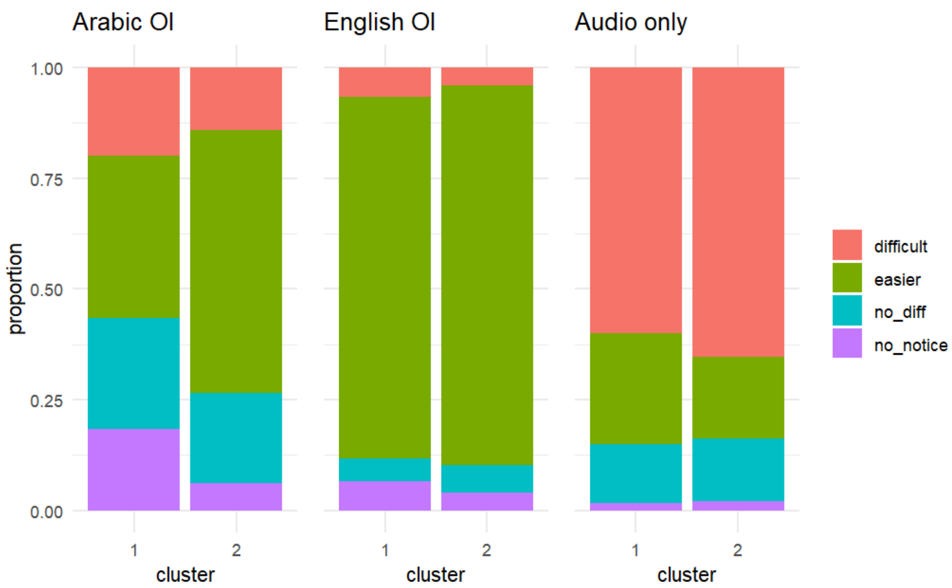
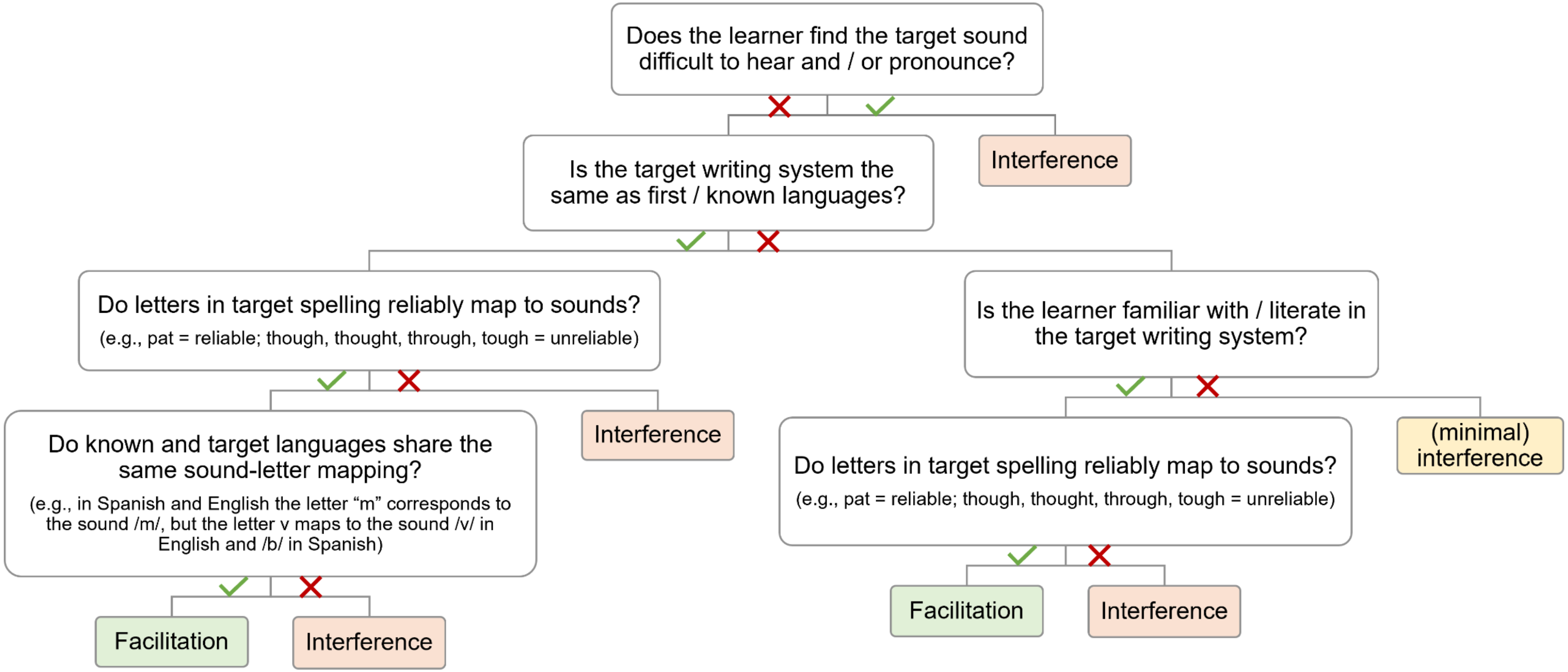


Figure XXII.5 Bar plots of perceived difficulty associated with OI conditions by cluster

Appendix XXIII: Teaching resource for assessing orthographic facilitation/interference



Appendix XXIV: Modified POLLS Inventory

The POLLS (Phonology and Orthography Language Learning Strategies) inventory presented below is a tool for both language teachers and learners to better understand the role of phonological and orthographic knowledge when learning new vocabulary. This inventory could also be used for research purposes; however, guidance provided here relates specifically to pedagogical opportunities for this kind of tool. The aim of the questions and statements below is twofold:

1. Responses offer reflective insight into learners' beliefs and strategies in relation to pronunciation and how written language influences learning.
2. The statements can provide ideas for alternative approaches and strategies when learning new sounds and words that may not have been considered.

The content of the inventory below is designed to provoke reflection and discussion of these topics, not to recommend a particular approach to teaching and learning. Indeed, responses are likely to reveal diverse individual beliefs and strategies. However, greater awareness of such beliefs and strategies offers opportunity for evaluation and adaptation of learning. This tool was developed as part of my PhD research, as well as being inspired by other published strategy inventories, focused on pronunciation and orthography (Berkil, 2008; Oxford, 1990; Pawlak, 2010; Peterson, 1997; Rokoszewska, 2012; Shen, 2005)

Suggested lesson integration:

The four open questions could be used in whole class or small group discussion to introduce the topic, ideally after the introduction of new vocabulary. The strategy statements would then be set as an individual task, before comparing different responses in small group discussions. Whole group discussion could then provide an opportunity for the teacher to draw together ideas for future language presentation and allow for reflection on teaching practice, as well as individual learning inside and outside the classroom. Finally, learners could choose 2-3 strategies to try out in class next time new vocabulary is introduced. The aim of these activities is to raise awareness and encourage evaluative thinking.

A. Open response questions:

1) What do you do when you are learning new words in your second/new language?
2) How do you learn the pronunciation of new words?
3) How important is accurate pronunciation of new words to you?
4) How does seeing the written form influence the learning of new words?

B. Multiple-choice questions:

Please read the statements below and choose a response from 1 to 5 which best describes your situation. The options are as follows:

1. Never
2. Rarely
3. Sometimes
4. Often
5. Always

B.1 When a new word is **first** introduced…

1) I focus on listening carefully to the pronunciation of the word.
2) I ask to see it written down or I look for the spelling.
3) I remember the location of the spelling/transcription on the page/screen/board
4) I use the pronunciation of my first/known languages, if the new word is spelled similarly.
5) I ignore/avoid the spelling, if it is presented together with the pronunciation.
6) I pay attention to difficult sounds in the new word.
7) I memorise the pronunciation first, then the spelling.
8) I memorise the spelling and the pronunciation together.
9) I use the spelling to help connect the sound and meaning of the word.

B.2 I use my experience reading and writing in my **first/known language(s)** to…

10) Feel more confident about the new word.
11) Visualise the spelling of the word in my mind.

12) Distinguish between similar sounding words.
13) Associate (difficult) sounds with specific letters or symbols.
14) Make a mental image connecting the spelling, sounds and meaning of a word.
15) Remember how to pronounce words with my own code/spelling/transcription.
16) Make/annotate notes about the pronunciation of the word.

B.3 In order to **remember** the word…

17) I create associations with words or things I already know.
18) I put the word in a context to remember it (e.g. a sentence, a story, a rhyme).
19) I associate the word with a mental image or draw a picture.
20) I think about an action or movement to help remember the words.
21) I group similar sounding words.
22) I place the new word in a group with other words that are similar in some way.
23) I connect individual sounds and letters.
24) I find patterns in the new words or sounds.
25) I break words down into syllables or sounds.
26) I connect specific letters to the image/meaning.
27) I remember the first sounds or letters.
28) I look for the small differences between words.
29) I look for similarities and contrasts between the new language and the language(s) I know.
30) I try to identify and use pronunciation rules.

B.4 In order to **practise the pronunciation** of new words...

31) I repeat the words out loud or silently in my head.
32) I practise difficult words or sounds over and over.
33) I think about or practise mouth positions to pronounce the words.
34) I whisper the word to focus on the feeling of articulation.
35) I notice the mouth positions/watch lips of my teacher/native speakers.
36) I imitate the pronunciation of my teacher/native speakers.
37) I guess pronunciation on the basis of spelling.

38) I say the word out loud or in my head when writing it.

39) I concentrate on my pronunciation when reading a word out loud.

40) I concentrate on a speaker’s pronunciation while listening.

B.5 In order to **improve my understanding and use** of new words…

41) I decide to focus on particular sounds.

42) I look for opportunities to practise new words or sounds.

43) I purposely avoid producing sounds from my other known languages.

44) I notice my pronunciation errors and find reasons for them.

45) I notice my mistakes and use that information to help me improve.

46) I test myself during learning/memorisation to assess my progress.

47) I reflect on previous learning to direct my attention and anticipate points of difficulty.

48) I find ways to test my memory and recall the new words.

49) I think about my progress in learning new words.

50) I have clear goals for improving my pronunciation.

Glossary

< >	Orthographic representation
/ /	Phonological representation
[ ]	Phonetic realisation
α	Alpha reliability coefficient
β	Beta regression coefficient
χ^2	Chi-square statistic
AXB	Perceptual discrimination task, where X matches either A or B.
BIAM	Bi-modal interactive-activation model
c	Criterion scores
CG	Category-goodness assimilation
CI	95% confidence interval
d′	d-prime scores
EFA	Exploratory factor analysis
GLMM	Generalised linear mixed-effects model
GPC	Grapheme-phoneme correspondence
ID	Individual differences
KMO	Kaiser-Meyer-Olkin statistic
L1	First language
L2	Second language
L2LP	Second language perception model
LP	Learner profile
LQH	Lexical quality hypothesis
LLS	Language learning strategies
M	mean
ms	milliseconds
MSA	Modern Standard Arabic
n	number of participants/observations
ODH	Orthographic depth hypothesis
OI	Orthographic input

p	p-value (probability of significant difference)
PAM	Perceptual assimilation model
PCA	Principal component analysis
PLS	Pronunciation learning strategies
POLLS	Phonology and orthography language learning strategies
QCA	Quantitative content analysis
r	Correlation coefficient
Rmd	R Markdown file
RT	Response time
S2R	Strategic self-regulation (model of language learning)
SC	Single-category assimilation
SD	Standard deviation
SE	Standard error
SILL	Strategy inventory for language learning
SLA	Second language acquisition
SLM	Speech learning model
SSBE	Southern Standard British English
t	t-value (size of difference relative to variation in data)
TC	Two-category assimilation
UCM	Unified competition model
WM	Working memory
z	z-score (standardised score)

References

Aboudan, R., Eapen, V., Bayshak, M., Al-Mansouri, M., & Al-Shamsi, M. (2011). Dyslexia in the United Arab Emirates university – A study of prevalence in English and Arabic. *International Journal of English Linguistics*, *1*(2). https://doi.org/10.5539/ijel.v1n2p64

Al Azmi, S. (2019). *Aural and orthographic input: Implications for the acquisition of English consonant clusters by Northern Najdi Arabic speakers* [Doctoral dissertation, Newcastle University]. Newcastle University theses repository. https://theses.ncl.ac.uk/jspui/handle/10443/4845

Albirini, A. (2018). The acquisition of Arabic as a first language. In E. Benmamoun & R. Bassiouney (Eds.), *Routledge Handbook of Arabic Linguistics* (pp. 227–248). Routledge.

Allan, D. (2004). *Oxford Placement Tests 1: Test Pack*. OUP Oxford.

Allehaiby, W. H. (2013). Arabizi: An analysis of the romanization of the Arabic script from a sociolinguistic perspective. *Arab World English Journal*, *4*(3), 52–62.

Al-Qadri, A. H., Zhao, W., Li, M., Al-Khresheh, M. H., & Boudouaia, A. (2021). The prevalence of academic learning difficulties: An observation tool. *Heliyon*, *7*(10), Article e08164.

Altaha, F. M. (1995). Pronunciation errors made by Saudi university students learning English. *ITL - International Journal of Applied Linguistics*, *109*(1), 110–123.

Amerstorfer, C. M. (2018). Past its expiry date? The SILL in modern mixed-methods strategy research. *Studies in Second Language Learning and Teaching*, *8*(2), 497–523.

Anderson, J. (1982). Acquisition of cognitive skill. *Psychological Review*, *89*(4), 369–406.

Anderson, J., Betts, S., Bothell, D., Hope, R. M., & Lebiere, C. (2018). Three aspects of skill acquisition. In *PsyArXiv*. https://doi.org/10.31234/osf.io/rh6zt

Andringa, S. (2020). The emergence of awareness in uninstructed L2 learning: A visual world eye tracking study. *Second Language Research*, *36*(3), 335–357.

Andringa, S., & Godfroid, A. (2020). Sampling bias and the problem of generalizability in applied linguistics. *Annual Review of Applied Linguistics*, *40*, 134–142.

Andringa, S., & Rebuschat, P. (2015). New directions in the study of implicit and explicit learning: An Introduction. *Studies in Second Language Acquisition*, *37*(2), 185–196.

Anthony, J. L., & Francis, D. J. (2005). Development of phonological awareness. *Current Directions in Psychological Science*, *14*(5), 255–259.

Antin, J., & Shaw, A. (2012). Social desirability bias and self-reports of motivation: a study of amazon mechanical turk in the US and India. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2925–2934.

Anwyl-Irvine, A. L., Dalmaijer, E. S., Hodges, N., & Evershed, J. K. (2021). Realistic precision and accuracy of online experiment platforms, web browsers, and devices. *Behavior Research Methods*, *53*, 1407–1425.

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*, *52*(1), 388–407.

Aoyama, K., Flege, J. E., Guion, S. G., Akahane-Yamada, R., & Yamada, T. (2004). Perceived phonetic dissimilarity and L2 speech learning: the case of Japanese /r/ and English /l/ and /r/. *Journal of Phonetics*, *32*(2), 233–250.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412.

Baddeley, A. (2012). Working memory: theories, models, and controversies. *Annual Review of Psychology*, *63*, 1–29.

Baddeley, A. (2017). *Exploring working memory: Selected works of Alan Baddeley*. Routledge.

Bakker, I., Takashima, A., van Hell, J. G., Janzen, G., & McQueen, J. M. (2014). Competition from unseen or unheard novel words: Lexical consolidation across modalities. *Journal of Memory and Language*, *73*, 116–130.

Barcelos, A. M. F., & Kalaja, P. (2011). Introduction to beliefs about SLA revisited. *System*, *39*(3), 281–289.

Bassetti, B. (2007). Effects of hanyu pinyin on pronunciation in learners of Chinese as a foreign language. In A. Guder, X. Jiang, & Y. Wan (Eds.), *The Cognition, Learning and Teaching of Chinese Characters* (p. 552). Beijing Language and Culture University Press.

Bassetti, B. (2017). Orthography affects second language speech: Double letters and geminate production in english. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *43*(11), 1835–1842.

Bassetti, B., & Atkinson, N. (2015). Effects of orthographic forms on pronunciation in experienced instructed second language learners. *Applied Psycholinguistics*, *36*(01), 67–91.

Bassetti, B., Escudero, P., & Hayes-Harb, R. (2015). Second language phonology at the interface between acoustic and orthographic input. *Applied Psycholinguistics*, *36*, 1–6.

Bassetti, B., Mairano, P., Masterson, J., & Cerni, T. (2020). Effects of orthographic forms on second language speech production and phonological awareness, with consideration of speaker-level predictors. *Language Learning*, *70*(4), 1218–1256.

Bassetti, B., Soković-Perović, M., Mairano, P., & Cerni, T. (2018). Orthography-induced length contrasts in the second language phonological systems of L2 speakers of English: Evidence from minimal pairs. *Language and Speech*, *61*(4), 577–597.

Berkil, G. (2008). *A closer look at pronunciation learning strategies, L2 pronunciation proficiency and secondary variables influencing pronunciation ability* (J. Walters (ed.)) [MA Dissertation, Bilkent University]. Bilkent University Institutional Repository. http://repository.bilkent.edu.tr/handle/11693/14676

Best, C. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 171–204). York Press.

Best, C., & Tyler, M. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In M. J. Munro & O. S. Bohn (Eds.), *Second language speech learning: The role of language experience in speech perception and production*. John Benjamins Publishing Company.

Bhaskar, R. (2009). *Scientific Realism and Human Emancipation*. Routledge.

Bialystok, E. (1978). A theoretical model of second language learning. *Language Learning*, *28*(1), 69–83.

Bialystok, E. (2001). *Bilingualism in development: Language, literacy, and cognition*. Cambridge University Press.

Bigelow, M., & Schwarz, R. L. (2010). *Adult English language learners with limited literacy*. National Institute for Literacy.

Binturki, T. A. (2008). *Analysis of pronunciation errors of Saudi ESL learners* [MA dissertation, Southern Illinois University Carbondale]. KU ScholarWorks. https://kuscholarworks.ku.edu/handle/1808/17214

Birnbaum, M. H. (Ed.). (2000). *Psychological experiments on the internet*. Elsevier.

Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. In M. A. Gernsbacher, R. W. Pew, L. M. Hough, & J. R. Pomerantz (Eds.), *Psychology and the real world: Essays illustrating fundamental contributions to society* (Vol. 2, pp. 59–64). Worth Publishers.

Bjork, R. A., & Bjork, E. L. (2020). Desirable difficulties in theory and practice. *Journal of Applied Research in Memory and Cognition*, *9*(4), 475–479.

Bjorndahl, C. (2022). Voicing and frication at the phonetics-phonology interface: An acoustic study of Greek, Serbian, Russian, and English. *Journal of Phonetics*, *92*, Article e101136.

Block, D. (2003). *The social turn in second language acquisition*. Edinburgh University Press.

Bloom, B. S. (1964). *Taxonomy of educational objectives: The classification of educational objectives* (Vol. 2). McKay.

Bogulski, C. A., Bice, K., & Kroll, J. F. (2019). Bilingualism as a desirable difficulty: Advantages in word learning depend on regulation of the dominant language. *Bilingualism* , *22*(5), 1052–1067.

Bordag, D., Gor, K., & Opitz, A. (2022). Ontogenesis Model of the L2 Lexical Representation. *Bilingualism: Language and Cognition*, *25*(2), 185–201.

Boudelaa, S. (2014). Is the Arabic mental lexicon morpheme-based or stem-based? Implications for spoken and written word recognition. In E. Saiegh-Haddad & R. M. Joshi (Eds.), *Handbook of Arabic Literacy: Insights and Perspectives* (pp. 31–54). Springer.

Bovolenta, G., & Marsden, E. (2022). Prediction and error-based learning in L2 processing and acquisition: A conceptual review. *Studies in Second Language Acquisition*, *44*(5), 1384–1409.

Bowers, J. S., & Bowers, P. N. (2017). Beyond phonics: The case for teaching children the logic of the English spelling system. *Educational Psychologist*, *52*(2), 124–141.

Box, G. E. P. (1976). Science and statistics. *Journal of the American Statistical Association*, *71*(356), 791–799.

Braun, V., & Clarke, V. (2013). *Successful Qualitative Research: A Practical Guide for Beginners*. SAGE.

Braun, V., & Clarke, V. (2021). Can I use TA? Should I use TA? Should I not use TA? Comparing reflexive thematic analysis and other pattern-based qualitative analytic approaches. *Counselling and Psychotherapy Research*, *21*(1), 37–47.

Brehm, L., & Alday, P. M. (2022). Contrast coding choices in a decade of mixed models. *Journal of Memory and Language*, *125*, Article e104334.

Bridges, D., Pitiot, A., MacAskill, M. R., & Peirce, J. W. (2020). The timing mega-study: comparing a range of experiment generators, both lab-based and online. *PeerJ*, *8*, Article e9414.

Brodeur, M. B., Dionne-Dostie, E., Montreuil, T., & Lepage, M. (2010). The Bank of Standardised Stimuli (BOSS): A new set of 480 normative photos of objects to be used as visual stimuli in cognitive research. *PloS One*, *5*(5), Article e10773.

Broersma, M. (2005). *Phonetic and lexical processing in a second language* [PhD dissertation, Radboud University Nijmegen]. Radboud Repository. https://repository.ubn.ru.nl/handle/2066/56388

Brown, G., & Hellmuth, S. (2022). Computational modelling of segmental and prosodic levels of analysis for capturing variation across Arabic dialects. *Speech Communication*, *141*, 80–92.

Brown, K. (2015). *The influence of explicit instruction on failure to acquire a phonological rule due to orthographic input: The case of native English speakers learning German* [MA dissertation, University of Utah]. University of Utah ProQuest.

Brown, V. A. (2021). An introduction to linear mixed-effects modeling in R. *Advances in Methods and Practices in Psychological Science*, *4*(1), 1–19.

Brucki, S. M. D., & Rocha, M. S. G. (2004). Category fluency test: effects of age, gender and education on total scores, clustering and switching in Brazilian Portuguese-speaking subjects. *Brazilian Journal of Medical and Biological Research*, *37*, 1771–1777.

Brysbaert, M., & Stevens, M. (2018). Power analysis and effect size in mixed effects models: A tutorial. *Journal of Cognition*, *1*(1), 9.

Bürki, A., Welby, P., Clément, M., & Spinelli, E. (2019). Orthography and second language word learning: Moving beyond “friend or foe?” *The Journal of the Acoustical Society of America*, *145*, 265.

Burnage, C. (2021, November 3). *Thematic Analysis: in conversation with Virginia Braun & Victoria Clarke*. Methodspace. https://www.methodspace.com/blog/thematic-analysis-in-conversation-with-virginia-braun-and-victoria-clarke

Całka, A. (2011). Pronunciation learning Strategies - Identification and classification. In Pawlak, M. Waniek-Klimczak, E. & Majer, J. (Ed.), *Speaking and instructed foreign language acquisition* (pp. 149–168). Multilingual Matters.

Carroll, S. E. (2013). Introduction to the special issue: Aspects of word learning on first exposure to a second language. *Second Language Research*, *29*(2), 131–144.

Cerni, T., Bassetti, B., & Masterson, J. (2019). Effects of orthographic forms on the acquisition of novel spoken words in a second language. *Frontiers in Communication*, *4*, Article 31.

Chamot, A. U. (2019). Language learning strategy instruction: Current issues and research. *Annual Review of Applied Linguistics*, *25*, 112–130.

Chéreau, C., Gaskell, M. G., & Dumay, N. (2007). Reading spoken words: Orthographic effects in auditory priming. *Cognition*, *102*(3), 341–360.

Child, D. (2006). *The essentials of factor analysis*. Continuum.

Cho, M. (2018). Task complexity and modality: Exploring learners' experience from the perspective of flow. *The Modern Language Journal*, *102*(1), 162–180.

Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review*, *3*(3), 149–210.

Clifford, S., & Jerit, J. (2014). Is there a cost to convenience? An experimental comparison of data quality in laboratory and online studies. *Journal of Experimental Political Science*, *1*(2), 120–131.

Colantoni, L., & Steele, J. (2008). Integrating articulatory constraints into models of second language phonological acquisition. *Applied Psycholinguistics*, *29*(3), 489–534.

Colantoni, L., Steele, J., & Escudero, P. (2015). *Second language speech: Theory and practice*. Cambridge University Press.

Coltheart, M. (2006). Dual route and connectionist models of reading: An overview. *London Review of Education*, *4*(1), 5–17.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*(1), 204–256.

Cook, S. V., Pandža, N. B., Lancaster, A. K., & Gor, K. (2016). Fuzzy nonnative phonolexical representations lead to fuzzy form-to-meaning mappings. *Frontiers in Psychology*, *7*, Article 1345.

Corder, S. P. (1967). The significance of learner's errors. *IRAL, International Review of Applied Linguistics in Language Teaching*, *5*(4), 161–170.

Crowther, D., Kim, S., Lee, J., Lim, J., & Loewen, S. (2021). Methodological synthesis of cluster analysis in second language research. *Language Learning*, *71*(1), 99–130.

Curcic, M., Andringa, S., & Kuiken, F. (2019). The role of awareness and cognitive aptitudes in L2 predictive language processing. *Language Learning*, *69*, 42–71.

Cutler, A. (2015). Representation of second language phonology. *Applied Psycholinguistics*, *36*(1), 115–128.

Darcy, I. (2017). *A psycholinguist walks into a classroom…': A road-map for bridging research and practice in pronunciation teaching (Plenary Talk)*. PSLLT, Salt Lake City.

Darcy, I., Daidone, D., & Kojima, C. (2013). Asymmetric lexical access and fuzzy lexical representations in second language learners. *The Mental Lexicon*, *8*(3), 372–420.

DeBruine, L., & Barr, D. J. (2019). Understanding mixed effects models through data simulation. *Advances in Methods and Practices in Psychological Science*, *4*(1), Article 2515245920965119.

Dehaene, S., Cohen, L., Morais, J., & Kolinsky, R. (2015). Illiterate to literate: behavioural and cerebral changes induced by reading acquisition. *Nature Reviews. Neuroscience*, *16*(4), 234–244.

DeKeyser, R. (2020). Skill acquisition theory. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (2nd ed., pp. 83–104). Routledge.

de Leeuw, J. R. (2015). jsPsych: a JavaScript library for creating behavioral experiments in a Web browser. *Behavior Research Methods*, *47*(1), 1–12.

Diependaele, K., Ziegler, J. C., & Grainger, J. (2010). Fast phonology and the bimodal interactive activation model. *The European Journal of Cognitive Psychology*, *22*(5), 764–778.

Dóczi, B. (2019). An overview of conceptual models and theories of lexical representation in the mental lexicon. In *The Routledge Handbook of Vocabulary Studies* (1st Ed, pp. 46–65). Routledge.

Dolnicar, S. (2002). A review of unquestioned standards in using cluster analysis for data-driven market segmentation. *The Australian and New Zealand Marketing Academy Conference (ANZMAC)*. https://ro.uow.edu.au/commpapers/273/

Dornyei, Z. (2005). *The psychology of the language learner: Individual differences in second* *language acquisition*. Routledge.

Dörnyei, Z., & Dewaele, J.-M. (2022). *Questionnaires in second language research: Construction, administration, and processing*. Routledge.

Dornyei, Z., & Ryan, S. (2015). *The psychology of the language learner revisited*. Routledge.

Dörnyei, Z., & Skehan, P. (2003). Individual differences in second language learning. *The Handbook of Second Language Acquisition*, *589*. https://onlinelibrary.wiley.com/doi/pdf/10.1002/9780470756492

do Val Barros, A. M. (2003). *Pronunciation difficulties in the consonant system experienced by Arabic speakers when learning English after the age of puberty* [Master's dissertation, West Virginia University]. WVU research repository. https://doi.org/10.33915/etd.766

Dumay, N., & Gaskell, G. (2012). Overnight lexical consolidation revealed by speech segmentation. *Cognition*, *123*(1), 119–132.

Dumay, N., & Gaskell, M. G. (2007). Sleep-associated changes in the mental representation of spoken words. *Psychological Science*, *18*(1), 35–39.

Dunteman, G. H. (1989). *Principal Components Analysis*. SAGE.

Elbro, C., & Pallesen, B. R. (2002). The quality of phonological representations and phonological awareness: A causal link. *Precursors of Functional Literacy*, *11*, 17–31.

Ellis, N. C. (2006). Cognitive perspectives on SLA: The associative-cognitive CREED. *AILA Review*, *19*(1), 100–121.

Ellis, N. C. (2008). Usage-based and form-focused language acquisition: The associative learning of constructions, learned attention, and the limited L2 endstate. In P. Robinson & E. N. C. (Eds.), *Handbook of cognitive linguistics and second language acquisition* (pp. 372–405). Routledge.

Ellis, N. C. (2015). Implicit AND Explicit Language Learning: Their dynamic interface and complexity. In P. Rebuschat (Ed.), *Implicit and explicit learning of languages* (pp. 3–23). John Benjamins.

Ellis, N. C., & Collins, L. (2009). Input and second language acquisition: The roles of frequency, form, and function introduction to the special issue. *Modern Language Journal*, *93*(3), 329–335.

Ellis, N. C., & Wulff, S. (2014). Usage-based approaches to SLA. In *Theories in second language acquisition* (pp. 87–105). Routledge.

Ellis, R. (2002). Does form-focused instruction affect the acquisition of implicit knowledge?: A review of the research. *Studies in Second Language Acquisition*, *24*(2), 223–236.

Erdener, V. D., & Burnham, D. K. (2005). The role of audiovisual speech and orthographic information in nonnative speech production. *Language Learning*, *55*(2), 191–228.

Escudero, P. (2005). *Linguistic perception and second language acquisition: Explaining the attainment of optimal phonological categorization*. LOT.

Escudero, P. (2015). Orthography plays a limited role when learning the phonological forms of new words: The case of Spanish and English learners of novel Dutch words. *Applied Psycholinguistics*, *36*(1), 7–22.

Escudero, P., & Boersma, P. (2004). Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition*, *26*(4), 551–585.

Escudero, P., & Hayes-Harb, R. (2022). The Ontogenesis Model may provide a useful guiding framework, but lacks explanatory power for the nature and development of L2 lexical representation. *Bilingualism: Language and Cognition*, *25*(2), 212–213.

Escudero, P., Hayes-Harb, R., & Mitterer, H. (2008). Novel second-language words and asymmetric lexical access. *Journal of Phonetics*, *36*(2), 345–360.

Escudero, P., Simon, E., & Mulak, K. E. (2014). Learning words in a new language: Orthography doesn't always help. *Bilingualism: Language and Cognition*, *17*(2), 384–395.

Escudero, P., Smit, E. A., & Angwin, A. J. (2022). Investigating orthographic versus auditory cross-situational word learning with online and laboratory-based testing. *Language Learning*, *73*(2), 543–577.

Eskenazi, M. A., & Nix, B. (2021). Individual differences in the desirable difficulty effect during lexical acquisition. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *47*(1), 45–52.

Evans, B. G., & Alshangiti, W. (2018). The perception and production of British English vowels and consonants by Arabic learners of English. *Journal of Phonetics*, *68*, 15–31.

Eviatar, Z., & Ibrahim, R. (2014). Why is it hard to read Arabic? In E. Saiegh-Haddad & R. M. Joshi (Eds.), *Handbook of Arabic literacy: Insights and perspectives* (pp. 77–96). Springer Netherlands.

Ferguson, C. A. (1959). Diglossia. *Word & World*, *15*(2), 325–340.

Field, A., Miles, J., & Field, Z. (2012). *Discovering statistics using R*. Sage publications.

Finn, A. S., Kalra, P. B., Goetz, C., Leonard, J. A., Sheridan, M. A., & Gabrieli, J. D. E. (2016). Developmental dissociation between the maturation of procedural memory and declarative memory. *Journal of Experimental Child Psychology*, *142*, 212–220.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, *92*, 233–277.

Flege, J. E., & Bohn, O.-S. (2021). The revised speech learning model (SLM-r). *Second Language Speech Learning: Theoretical and Empirical Progress*, 3–83.

Froud, K., & Khamis-Dakwar, R. (2021). The study of Arabic language acquisition: A critical review. In K. Ryding & D. Wilmsen (Eds.), *The Cambridge handbook of Arabic linguistics*. Cambridge University Press.

Furmanski, C. S., & Engel, S. A. (2000). Perceptual learning in object recognition: object specificity and size invariance. *Vision Research*, *40*(5), 473–484.

Galesic, M. (2006). Dropouts on the web: Effects of interest and burden experienced during an online survey. *Journal of Official Statistics*, *22*(2), 313–328.

Gaskell, M. G., & Marslen-Wilson, W. D. (1997). Integrating form and meaning: A distributed model of speech perception. *Language and Cognitive Processes*, *12*(5-6), 613–656.

Gaskell, M. G., Warker, J., Lindsay, S., Frost, R., Guest, J., Snowdon, R., & Stackhouse, A. (2014). Sleep underpins the plasticity of language production. *Psychological Science*, *25*(7), 1457–1465.

Gass, S. M., & Mackey, A. (2006). Input, interaction and output: An overview. *AILA Review*, *19*(1), 3–17.

Gates, G. A., & Mills, J. H. (2005). Presbycusis. *The Lancet*, *366*(9491), 1111–1120.

Gee, J. (1991). Socio-cultural approaches to literacy (literacies). *Annual Review of Applied Linguistics*, *12*, 31–48.

Gee, J. P. (2001). Reading as situated language: A sociocognitive perspective. *Journal of Adolescent & Adult Literacy: A Journal from the International Reading Association*, *44*(8), 714–725.

Gläser-Zikuda, M., Hagenauer, G., & Stephan, M. (2020). The potential of qualitative content analysis for empirical educational research. *Forum Qualitative Sozialforschung/ Forum: Qualitative Social Research*, *21*(1), Article 17.

Godfroid, A., Boers, F., & Housen, A. (2013). An eye for words: Gauging the role of attention in incidental L2 vocabulary acquisition by means of eye-tracking. *Studies in Second Language Acquisition*, *35*(3), 483–517.

Goody, J. (1987). *The Interface Between the Written and the Oral*. Cambridge University Press.

Gorilla Experiment Builder. (n.d.). Retrieved March 8, 2021, from https://gorilla.sc/support/info/timing

Gor, K., Cook, S., Bordag, D., Chrabaszcz, A., & Opitz, A. (2021). Fuzzy lexical representations in adult second language speakers. *Frontiers in Psychology*, *12*, Article 732030.

Goswami, U. (2002). Phonology, reading development, and dyslexia: A cross-linguistic perspective. *Annals of Dyslexia*, *52*(1), 139–163.

Goswami, U., & Bryant, P. (2016). *Phonological skills and learning to read*. Psychology Press.

Gould, S. J. J., Cox, A. L., Brumby, D. P., & Wiseman, S. (2015). Home is where the lab is: A comparison of online and lab data from a time-sensitive study of interruption. *Human Computation*, *2*(1), 45–67.

Graham, J. M., Guthrie, A. C., & Thompson, B. (2003). Consequences of not interpreting structure coefficients in published CFA research: A reminder. *Structural Equation Modeling: A Multidisciplinary Journal*, *10*(1), 142–153.

Grainger, J., & Ferrand, L. (1994). Phonology and orthography in visual word recognition: Effects of masked homophone primes. *Journal of Memory and Language*, *33*(2), 218–233.

Grainger, J., & Ziegler, J. C. (2008). Cross-code consistency in a functional architecture for word recognition. *Single-Word Reading: Behavioral and Biological Perspectives.*, *429*, 129–157.

Grainger, J., & Ziegler, J. C. (2011). A dual-route approach to orthographic processing. *Frontiers in Psychology*, *2*, 54.

Griffiths, C. (2018). *The strategy factor in successful language learning: The tornado effect* (2nd Ed). Multilingual Matters.

Guediche, S., Navarra-Barindelli, E., & Martin, C. D. (2023). Noise modulates crosslinguistic effects on second language auditory word recognition. *Journal of Speech, Language, and Hearing Research: JSLHR*, *66*(2), 635–647.

Guion, S. G., Clark, J. J., Harada, T., & Wayland, R. P. (2003). Factors affecting stress placement for English nonwords include syllabic structure, lexical class, and stress patterns of phonologically similar words. *Language and Speech*, *46*(Pt 4), 403–427.

Haeri, N. (2000). Form and ideology: Arabic sociolinguistics and beyond. *Annual Review of Anthropology*, *29*, 61–87.

Hall, N. (2017). Articulatory phonology. In S. J. Hannahs & A. Bosch (Eds.), *The Routledge handbook of phonological theory* (pp. 530–552). Routledge.

Han, Z., & Liu, Z. (2013). Input processing of Chinese by ab initio learners. *Second Language Research*, *29*(2), 145–164.

Hao, Y.-C., & Yang, C.-L. M. (2021). The effect of second-language orthographic input on the phonological encoding of Mandarin words. *Applied Psycholinguistics*, *42*, 887–906.

Hassan, E. M. I. (2014). Pronunciation problems: A case study of English language students at Sudan University of Science and Technology. *English Language and Literature Studies*, *4*(4), 31–44.

Hauser, D. J., & Schwarz, N. (2016). Attentive Turkers: MTurk participants perform better on online attention checks than do subject pool participants. *Behavior Research Methods*, *48*(1), 400–407.

Hautus, M. J. (1995). Corrections for extreme proportions and their biasing effects on estimated values of d′. *Behavior Research Methods, Instruments, & Computers*, *27*(1), 46–51.

Hayes, A. F., & Krippendorff, K. (2007). Answering the Call for a Standard Reliability Measure for Coding Data. *Communication Methods and Measures*, *1*(1), 77–89.

Hayes-Harb, R. (2006). Native speakers of Arabic and ESL texts: Evidence for the transfer of written word identification processes. *TESOL Quarterly*, *40*(2), 321.

Hayes-Harb, R., & Barrios, S. (2021). The influence of orthography in second language phonological acquisition. *Language Teaching*, 1–30.

Hayes-Harb, R., Brown, K., & Smith, B. L. (2018). Orthographic input and the acquisition of german final devoicing by native speakers of English. *Language and Speech*, *61*(4), 547–564.

Hayes-Harb, R., & Cheng, H. W. (2016). The influence of the Pinyin and Zhuyin writing systems on the acquisition of Mandarin word forms by native English speakers. *Frontiers in Psychology*, *7*, Article 785.

Hayes-Harb, R., & Hacking, J. (2015). The Influence of written stress marks on native English speakers' acquisition of Russian lexical stress contrasts. *The Slavic and East European Journal*, *59*(1), 91–109.

Hayes-Harb, R., Nicol, J., & Barker, J. (2010). Learning the phonological forms of new words: effects of orthographic and auditory input. *Language and Speech*, *53*(3), 367–381.

Hellmuth, S. (2013). Phonology. In J. Owens (Ed.), *The Oxford handbook of Arabic Linguistics* (pp. 45–70). Oxford University Press.

Hoppe, D. B., Hendriks, P., Ramscar, M., & van Rij, J. (2022). An exploration of error-driven learning in simple two-layer networks from a discriminative learning perspective. *Behavior Research Methods*, *54*(5), 2221–2251.

Horlyck, S., Reid, A., & Burnham, D. (2012). The relationship between learning to read and language-specific speech perception: Maturation versus experience. *Scientific Studies of Reading*, *16*(3), 218–239.

Horst, J. S., & Hout, M. C. (2016). The Novel Object and Unusual Name (NOUN) Database: A collection of novel images for use in experimental research. *Behavioural Research Methods*, *48*(4), 1393–1409.

Huettig, F., Kolinsky, R., & Lachmann, T. (2018). The culturally co-opted brain: how literacy affects the human mind. *Language, Cognition and Neuroscience*, *33*(3), 275–277.

Huettig, F., & Mishra, R. K. (2014). How literacy acquisition affects the illiterate mind - A critical examination of theories and evidence. *Language and Linguistics Compass*, *8*(10), 401–427.

Hulme, C. (2002). Phonemes, rimes, and the mechanisms of early reading development. *Journal of Experimental Child Psychology*, *82*(1), 58–64.

Hulstijn, J. H. (2005). Theoretical and empirical issues in the study of implicit and explicit second language learning: An introduction. *Studies in Second Language Acquisition*, *27*(2), 129–140.

Hulstijn, J. H., Young, R. F., Ortega, L., Bigelow, M., Dekeyser, R., Mackey, A., & Talmy, S. (2014). Bridging the gap: Cognitive and social approaches to research in second language learning and teaching. *Studies in Second Language Acquisition*, *36*, 361–421.

Izquierdo, I., Olea, J., & Abad, F. J. (2014). Exploratory factor analysis in validation studies: Uses and recommendations. *Psicothema*, *26*(3), 395–400.

Jackson, J. L. (2016). *The effects of novel orthographic elements and phonetic instruction in second* *language phonological acquisition* [Master's dissertation, University of Utah]. Digital Library The University of Utah. https://collections.lib.utah.edu/ark:/87278/s6gn1gn3

Jacques, J., & Biernacki, C. (2018). Model-based co-clustering for ordinal data. *Computational Statistics & Data Analysis*, *123*, 101–115.

Jarvis, S., & Pavlenko, A. (2008). *Crosslinguistic Influence in Language and Cognition*. Routledge.

Jongman, A., Wayland, R., & Wong, S. (2000). Acoustic characteristics of English fricatives. *The Journal of the Acoustical Society of America*, *108*(3), 1252–1263.

Jun, E., Hsieh, G., & Reinecke, K. (2017). Types of motivation affect study selection, attention, and dropouts in online experiments. *Proceedings of the ACM on Human-Computer Interaction*, *1*(CSCW), Article 56.

Kachru, B. B. (1997). World Englishes and English-using communities. *Annual Review of Applied Linguistics*, *17*, 66–87.

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, *39*(1), 31–36.

Kapnoula, E. C., & McMurray, B. (2016). Newly learned word forms are abstract and integrated immediately after acquisition. *Psychonomic Bulletin & Review*, *23*(2), 491–499.

Katz, L., & Frost, R. (1992). The reading process is different for different orthographies: The orthographic depth hypothesis. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology, and meaning* (pp. 67–84). Elsevier.

Kaufmann, N., Schulze, T., & Veit, D. (2011). More than fun and money. Worker motivation in crowdsourcing - A study on mechanical turk. *The Seventeenth Americas Conference on Information Systems*.

Kharma, N., & Hajjaj, A. (1989). *Errors in English among Arabic speakers: Analysis and remedy*. Longman.

Koda, K. (2007). Phonology and literacy. In M. C. Pennington (Ed.), *Phonology in Context* (pp. 219–244). Palgrave Macmillan UK.

Koda, K. (2013). Second language reading, scripts, and orthographies. In C. A. Chapelle (Ed.), *The encyclopaedia of applied linguistics* (pp. 1–7). Blackwell Publishing.

Krashen, S. D. (1977). The monitor model for adult second language performance. In M. Burt, H. Dulay, & M. Finocchiaro (Eds.), *Viewpoints on English as a second language* (pp. 152–161). Regents.

Krashen, S. D. (1981). *Second language acquisition and second language learning*. Pergamon.

Krashen, S. D. (1985). *The input hypothesis: Issues and implications*. Longman.

Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory into Practice*, *41*(4), 212–218.

Krippendorff, K. (2004). Reliability in content analysis. *Human Communication Research*, *30*(3), 411–433.

Krippendorff, K. (2017). Intercoder reliability techniques: Krippendorff's Alpha. In M. Allen (Ed.), *The SAGE encyclopedia of communication research methods* (pp. 744–751). Sage Publications.

Kuhl, P. K., Conboy, B. T., Padden, D., Nelson, T., & Pruitt, J. (2005). Early speech perception and later language development: Implications for the "critical period." *Language Learning and Development: The Official Journal of the Society for Language Development*, *1*(3-4), 237–264.

Lange, K., Kühn, S., & Filevich, E. (2015). "Just Another Tool for Online Studies" (JATOS): An easy solution for setup and management of web servers supporting online studies. *PloS One*, *10*(6), Article e0130834.

Lethaby, C., & Mayne, R. (2020). A critical examination of perceptual learning styles in English language teaching. *IRAL, International Review of Applied Linguistics in Language Teaching*, *58*(2), 221–237.

Li, S., Hiver, P., & Papi, M. (2022). *The Routledge handbook of second language acquisition and individual differences*. Routledge.

Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413–468). Elsevier.

Lo, S., & Andrews, S. (2015). To transform or not to transform: Using generalized linear mixed models to analyse reaction time data. *Frontiers in Psychology*, *6*, Article 1171.

Luna, F. (2009). Elucidating the concept of vulnerability: Layers not labels. *IJFAB: International Journal of Feminist Approaches to Bioethics*, *2*(1), 121–139.

Luna, F. (2019). Identifying and evaluating layers of vulnerability - a way forward. *Developing World Bioethics*, *19*(2), 86–95.

MacWhinney, B. (1987). The competition model. In B. MacWhinney (Ed.), *Mechanisms of* *language acquisition: The 20th Annual Carnegie Mellon symposium on cognition* (pp. 249–308).

MacWhinney, B. (2002). The competition model: The input, the context, and the brain. In P. Robinson (Ed.), *Cognition and second language instruction*. Cambridge University Press.

MacWhinney, B. (2008). A unified model. In N. C. Ellis & P. Robinson (Eds.), *Handbook of cognitive linguistics and second language acquisition* (pp. 341–371). Routledge.

MacWhinney, B. (2012). The logic of the unified model. In S. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 229–245). Routledge.

MacWhinney, B. (2018). A unified model of first and second language learning. In M. Mickmann, E. Veneziano, & H. Jisa (Eds.), *Sources of variation in first language acquisition: Languages, contexts and learners* (pp. 287–312). John Benjamins.

Maddieson, I. (1984). *Patterns of Sounds*. Cambridge University Press.

Major, R. (1987). The natural phonology of second language acquisition. In A. James & J. Leather (Eds.), *Sound patterns in second language acquisition* (pp. 207–224). De Gruyter Mouton.

Manly, J. J., Jacobs, D. M., Sano, M., Bell, K., Merchant, C. A., Small, S. A., & Stern, Y. (1999). Effect of literacy on neuropsychological test performance in nondemented, education-matched elders. *Journal of the International Neuropsychological Society: JINS*, *5*(3), 191–202.

Mansur, L. L., Radanovic, M., Araújo, G. de C., Taquemori, L. Y., & Greco, L. L. (2006). Boston Naming Test: Performance of Brazilian population from São Paulo. *Pro-Fono: Revista de Atualizacao Cientifica*, *18*(1), 13–20.

Marsden, E., Alferink, I., Andringa, S., & Bolibaugh, C. (2018). *Open Accessible Summaries in Language Studies (OASIS)*. OASIS Database. https://www.oasis-database.org

Mason, W., & Suri, S. (2012). Conducting behavioral research on Amazon's Mechanical Turk. *Behavior Research Methods*, *44*(1), 1–23.

Mathieu, L. (2016). The influence of foreign scripts on the acquisition of a second language phonological contrast. *Second Language Research*, *32*(2), 145–170.

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: an open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, *44*(2), 314–324.

Mathuranath, P. S., George, A., Cherian, P. J., Alexander, A., Sarma, S. G., & Sarma, P. S.

(2003). Effects of age, education and gender on verbal fluency. *Journal of Clinical and Experimental Neuropsychology*, *25*(8), 1057–1064.

Matusheck, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language*, *94*, 305–315.

May, L., Byers-Heinlein, K., Gervain, J., & Werker, J. F. (2011). Language and the newborn brain: does prenatal language experience shape the neonate neural response to speech? *Frontiers in Psychology*, *2*, 222.

Mayring. (2019). Qualitative content analysis: Demarcation, varieties, developments. *Forum, Qualitative Social Research*, *20*(3), Article 16.

Mayring, P. (2014). *Qualitative content analysis: theoretical foundation, basic procedures and software solution*. Klagenfurt.

Mayring, P. (2015). Qualitative content analysis: Theoretical background and procedures. In A. Bikner-Ahsbahs, C. Knipping, & N. Presmeg (Eds.), *Approaches to qualitative research in mathematics education: Advances in mathematics education* (pp. 365–380). Springer.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*(1), 1–86.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. an account of basic findings. *Psychological Review*, *88*(5), 375–407.

McLaughlin. (1987). *Theories of second language learning*. Edward Arnold.

Mejdell, G. (2018). Diglossia. In E. Benmamoun & R. Bassiouney (Eds.), *The Routledge handbook of Arabic linguistics* (pp. 332–344). Routledge.

Miller, R., Schmidt, K., Kirschbaum, C., & Enge, S. (2018). Comparability, stability, and reliability of internet-based mental chronometry in domestic and laboratory settings. *Behavior Research Methods*, *50*(4), 1345–1358.

Mitchell, R., Myles, F., & Marsden, E. (2013). *Second Language Learning Theories* (3rd Edition). Routledge.

Mok, P. P. K., Lee, A., Li, J. J., & Xu, R. B. (2018). Orthographic effects on the perception and production of L2 mandarin tones. *Speech Communication*, *101*, 1–10.

Morais, J., Bertelson, P., Cary, L., & Alegria, J. (1986). Literacy training and speech segmentation. *Cognition*, *24*(1-2), 45–64.

Morais, J., Cary, L., Alegria, J., & Bertelson, P. (1979). Does awareness of speech as a sequence of phones arise spontaneously? *Cognition*, *7*(4), 323–331.

Morgan-Short, K., Faretta-Stutenberg, M., Brill-Schuetz, K. A., Carpenter, H., & Wong, P. C. M. (2014). Declarative and procedural memory as individual differences in second language acquisition. *Bilingualism: Language and Cognition*, *17*(1), 56–72.

Mustafawi, E. (2017). Arabic phonology. In *The Routledge handbook of Arabic linguistics* (pp. 11–31). Routledge.

Nagle, C. L., & Baese-Berk, M. M. (2022). Advancing the state of the art in L2 speech perception-production research: revisiting theoretical assumptions and methodological practices. *Studies in Second Language Acquisition*, *44*(2), 580–605.

Narayan, C. (2008). The acoustic–perceptual salience of nasal place contrasts. *Journal of Phonetics*, *36*(1), 191–217.

Narayan, C., Werker, J., & Beddor, P. (2010). The interaction between acoustic salience and language experience in developmental speech perception: evidence from nasal place discrimination. *Developmental Science*, *13*(3), 407–420.

Nation, K., & Snowling, M. J. (2004). Beyond phonological skills: broader language skills contribute to the development of reading. *Journal of Research in Reading*, *27*(4), 342–356.

Nelson, J. R., Balass, M., & Perfetti, C. A. (2005). Differences between written and spoken input in learning new words. *Written Language and Literacy*, *8*(2), 25–44.

Neuendorf, K. (2018). Content analysis and thematic analysis. In K. Neuendorf (Ed.), *Advanced research methods for applied psychology* (pp. 211–223). Routledge.

Nimz, K., & Khattab, G. (2020). On the role of orthography in L2 vowel production : The case of Polish learners of German. *Second Language Research*, *36*(4), 623–652.

O'Connell, D. C., & Kowal, S. (2003). Psycholinguistics: a half century of monologism. *The American Journal of Psychology*, *116*(2), 191–212.

O'Connor, C., & Joffe, H. (2020). Intercoder reliability in qualitative research: Debates and practical guidelines. *International Journal of Qualitative Methods*, *19*, 1–13.

O'Grady, W. (2005). *Syntactic carpentry: An emergentist approach to syntax*. Routledge.

O'Grady, W. (2008). Innateness, universal grammar, and emergentism. *Lingua. International Review of General Linguistics.*, *118*(4), 620–631.

Olson, D. R., & Torrance, N. (1991). *Literacy and orality*. Cambridge University Press.

O'Malley, J. M., & Chamot, A. U. (1990). *Learning strategies in second language acquisition*. Cambridge University Press.

O'Malley, J. M., Chamot, A. U., & Walker, C. (1987). Some applications of cognitive theory to second language acquisition. *Studies in Second Language Acquisition*, *9*(3), 287–306.

Ortega, L. (2011). SLA after the social turn. *Alternative Approaches to Second Language Acquisition*, 167–180.

Ota, M., Hartsuiker, R. J., & Haywood, S. L. (2009). The KEY to the ROCK: Near-homophony in nonnative visual word recognition. *Cognition*, *111*(2), 263–269.

Owens, J. (2013). *The Oxford handbook of Arabic linguistics*. OUP.

Oxford, R. L. (1990). *Language learning strategies: What every teacher should know*. Newbury House Publisher.

Oxford, R. L. (1996). Employing a questionnaire to assess the use of language learning strategies. *Applied Language Learning*, *7*(1), 25–45.

Oxford, R. L. (2011). *Teaching & researching: Language learning strategies*. Routledge.

Oxford, R. L. (2016). *Teaching and researching language learning strategies: Self-regulation in context*. Routledge.

Papadatou-Pastou, M., Touloumakos, A. K., Koutouveli, C., & Barrable, A. (2021). The learning styles neuromyth: When the same term means different things to different teachers. *European Journal of Psychology of Education*, *36*(2), 511–531.

Pawlak, M. (2010). Designing and piloting a tool for the measurement of the use of pronunciation learning strategies. *Research in Language*, *8*, 189–202.

Pawlak, M. (2021). Investigating language learning strategies: Prospects, pitfalls and challenges. *Language Teaching Research*, *25*(5), 817–835.

Pawlak, M., & Szyszka, M. (2018). Researching pronunciation learning strategies: An overview and a critical look. *Studies in Second Language Learning and Teaching*, *8*(2), 293–323.

Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1), 8–13.

Perfetti, C. A. (2017). Lexical quality revisited. *Developmental Perspectives in Written Language and Literacy: In Honor of Ludo Verhoeven*, 51–67.

Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. In *Studies in Written Language* *and Literacy* (pp. 189–213). https://doi.org/10.1075/swll.11.14per

Perre, L., & Ziegler, J. C. (2008). On-line activation of orthography in spoken word recognition. *Brain Research*, *1188*, 132–138.

Peterson, S. S. (1997). *Pronunciation learning strategies and learning strategies related to pronunciation ability in American university students studying Spanish*. [Doctoral dissertation, Ohio State University] https://etd.ohiolink.edu/apexprod/rws_etd/send_file/send?accession=osu1487945320757394

Peyton, K., Huber, G. A., & Coppock, A. (2022). The generalizability of online experiments conducted during the COVID-19 pandemic. *Journal of Experimental Political Science*, *9*, 379–394.

Pierrehumbert, J. B. (2016). Phonological representation: Beyond abstract versus episodic. *Annual Review of Linguistics*, *2*(1), 33–52.

Piske, T., & Young-Scholten, M. (2008). *Input matters in SLA*. Multilingual Matters.

Plonsky, L. (2023). Sampling and generalizability in Lx research: A second-order synthesis. *Languages*, *8*(1), 75.

Plonsky, L., & Gonulal, T. (2015). Methodological synthesis in quantitative L2 research: A review of reviews and a case study of exploratory factor analysis. *Language Learning*, *65*(1), 9–36.

Price, L., & Martin, L. (2018). Introduction to the special issue: applied critical realism in the social sciences. *Journal of Critical Realism*, *17*(2), 89–96.

Procházka, S. (2021). Arabic dialectology. In K. Ryding & D. Wilmsen (Eds.), *The Cambridge handbook of Arabic linguistics*. Cambridge University Press.

Pronk, T., Wiers, R. W., Molenkamp, B., & Murre, J. (2020). Mental chronometry in the pocket? Timing accuracy of web applications on touchscreen and keyboard devices. *Behavior Research Methods*, *52*(3), 1371–1382.

Quam, C., Wang, A., Maddox, W. T., Golisch, K., & Lotto, A. (2018). Procedural memory, working memory, and declarative memory skills are each associated with dimensional integration in sound category learning. *Frontiers in Psychology*, *9*, Article 1828.

Rafat, Y. (2016). Orthography-induced transfer in the production of English-speaking learners of Spanish. *The Language Learning Journal*, *44*(2), 197–213.

Rafat, Y., & Perry, S. J. (2019). Navigating orthographic issues in the teaching of Spanish pronunciation. In R. Rao (Ed.), *Key issues in teaching of Spanish pronunciation*. Routledge.

Rafat, Y., & Stevenson, R. A. (2018). Auditory-orthographic integration at the onset of L2 speech acquisition. *Language and Speech*, *62*(3), 427–451.

Rastle, K. (2019). The place of morphology in learning to read in English. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *116*, 45–54.

Reis, A., Faísca, L., Ingvar, M., & Petersson, K. M. (2006). Color makes a difference: Two-dimensional object naming in literate and illiterate subjects. *Brain and Cognition*, *60*, 49–54.

Reis, A., Guerreiro, M., & Castro-Caldas, A. (1994). Brief report influence of educational level of non brain-damaged subjects on visual naming capacities. *Journal of Clinical and Experimental Neuropsychology*, *16*(6), 939–942.

Reis, A., Petersson, K. M., Castro-Caldas, A., & Ingvar, M. (2001). Formal schooling influences two but not three dimensional naming skills. *Brain and Cognition*, *47*(3), 397–411.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non-reinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (Vol. 2, pp. 64–69). Appleton-Century-Crofts.

Roach, P. (2004). British English: Received Pronunciation. *Journal of the International Phonetic Association*, *34*(2), 239–245.

Roberts, L., & Siyanova-Chanturia, A. (2013). Using eye-tracking to investigate topics in L2 acquisition and L2 processing. *Studies in Second Language Acquisition*, *35*(2), 213–235.

Robinson, P., Mackey, A., Gass, S. M., & Schmidt, R. (2013). Attention and awareness in second language acquisition. In *The Routledge handbook of second language acquisition* (pp. 247–267). Routledge.

Rodd, J. (2019, February 27). *How to maintain data quality when you can't see your participants*. Association for Psychological Science. https://www.psychologicalscience.org/observer/how-to-maintain-data-quality-when-you-cant-see-your-participants

Rokoszewska, K. (2012). The influence of pronunciation learning strategies on mastering

English vowels. *Studies in Second Language Learning and Teaching*, *2*(3), 391.

Rosselli, M., Ardila, A., & Rosas, P. (1990). Neuropsychological assessment in illiterates. II. Language and praxic abilities. *Brain and Cognition*, *12*(2), 281–296.

Saiegh–Haddad, E. (2004). The impact of phonemic and lexical distance on the phonological analysis of words and pseudowords in a diglossic context. *Applied Psycholinguistics*, *25*(4), 495–512.

Saiegh-Haddad, E., & Spolsky, B. (2014). Acquiring literacy in a diglossic context: Problems and prospects. In E. Saiegh-Haddad & R. M. Joshi (Eds.), *Handbook of Arabic literacy: Insights and perspectives* (pp. 225–240). Springer.

Saigh, K., & Schmitt, N. (2012). Difficulties with vocabulary word form: The case of Arabic ESL learners. *System*, *40*(1), 24–36.

Sauter, M., Draschkow, D., & Mack, W. (2020). Building, hosting and recruiting: A brief introduction to running behavioral experiments online. *Brain Sciences*, *10*, Article 251.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, *11*(2), 129–158.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3–32). Cambridge University Press.

Seaman, M. (2011). Bloom's taxonomy. *Curriculum & Teaching Dialogue*, *13*.

Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics in Language Teaching*, *10*(1-4), 209–232.

Seymour, P. H. K., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology* , *94*(Pt 2), 143–174.

Shad, D., Vasishth, S., Hohenstein, S., & Reinhold, K. (2020). How to capitalize on a priori contrasts in linear (mixed) models: A tutorial. *Journal of Memory and Language*, *110*, Article 104038.

Shafak, E. (2020). *How to Stay Sane in an Age of Division: The powerful, pocket-sized manifesto*. Wellcome collection.

Shafiro, V., Levy, E. S., Khamis-Dakwar, R., & Kharkhurin, A. (2012). Perceptual confusions of American-English vowels and consonants by native Arabic bilinguals. *Language and Speech*, *56*(2), 145–161.

Share, D. L. (2008). On the Anglocentricities of current reading research and practice: the perils of overreliance on an "outlier" orthography. *Psychological Bulletin*, *134*(4), 584–615.

Share, D. L. (2021). Is the science of reading just the science of reading English? *Reading Research Quarterly*, *56*(1), 391–402.

Sharwood Smith, M. (1993). Input enhancement in instructed SLA: Theoretical bases. *Studies in Second Language Acquisition*, *15*(2), 165–179.

Shen, H. H. (2005). An investigation of Chinese-character learning strategies among non-native speakers of chinese. *System*, *33*(1), 49–68.

Shepperd, L. (2013). *Influence of orthographic input on phonological representations of nonnative nonassimilable contrasts: Adult native English-speakers acquisition of Zulu click consonants* [Unpublished master's dissertation, University of Edinburgh].

Shepperd, L. (2018). *Influence of written input and script familiarity on adult L2 early word learning: Native English speakers' acquisition of Arabic consonants* [Unpublished master's dissertation, University of York].

Shepperd, L. (2022). Including underrepresented language learners in SLA research: A case study and considerations for internet-based methods. *Research Methods in Applied Linguistics*, *1*(3), Article 100031.

Showalter, C. E. (2018). Impact of Cyrillic on native English speakers' phono-lexical acquisition of Russian. *Language and Speech*, *61*(4), 565–576.

Showalter, C. E. (2020). Russian phono-lexical acquisition and orthographic input: Naïve learners, experienced learners, and interventions. *Studies in Second Language Acquisition*, *42*(2), 255–277.

Showalter, C. E., & Hayes-Harb, R. (2013). Unfamiliar orthographic information and second language word learning: A novel lexicon study. *Second Language Research*, *29*(2), 185–200.

Showalter, C. E., & Hayes-Harb, R. (2015). Native English speakers learning Arabic: The influence of novel orthographic information on second language phonological acquisition. *Applied Psycholinguistics*, *36*(1), 23–42.

Silveira, R. (2012). Production of English word-final consonants: The role of orthography and learner profile variables. *Trabalhos Em Linguística Aplicada*, *51*(1), 13–34.

Skeide, M. A., Kumar, U., Mishra, R. K., Tripathi, V. N., Guleria, A., Singh, J. P., Eisner, F., & Huettig, F. (2017). Learning to read alters cortico-subcortical cross-talk in the visual system of illiterates. *Science Advances*, *3*(5), Article e1602612.

Smith, B. (2001). Arabic speakers. In S. M. S. Bernard (Ed.), *Learner English: A teacher's guide to interference and other problems* (2nd Ed, pp. 195–213). Cambridge University Press.

Soković-Perović, M., Bassetti, B., & Dillon, S. (2019). English orthographic forms affect L2 English speech production in native users of a non-alphabetic writing system. *Bilingualism: Language and Cognition*, 1–11.

Spolsky, B. (2003). Religion as a site of language contact. *Annual Review of Applied Linguistics*, *23*, 81–94.

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers: A Journal of the Psychonomic Society, Inc*, *31*(1), 137–149.

Staples, S., & Biber, D. (2015). Cluster analysis. In L. Plonsky (Ed.), *Advancing quantitative methods in second language research*. Routledge.

Steinberger, F. (2017). *Synchronous collaborative L2 writing with technology: Interaction and learning* [Doctoral dissertation, LMU München]. LMU elektronische hochschulschriften. https://doi.org/10.5282/edoc.21677

Stoehr, A., & Martin, C. D. (2022). The impact of orthographic forms on speech production and perception: An artificial vowel-learning study. *Journal of Phonetics*, *94*, Article 101180.

Strauss, A., & Corbin, J. M. (1990). *Basics of qualitative research: Grounded theory procedures and techniques*. Sage Publications.

Strauss, A. L. (1987). *Qualitative analysis for social scientists*. Cambridge University Press.

Street, B. (1984). *Literacy in theory and practice*. Cambridge University Press.

Street, B. (1994). The new literacy studies: Implications for education and pedagogy. *Changing English*, *1*(1), 113–126.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643.

Sung, K.-Y., & Wu, H.-P. (2011). Factors influencing the learning of Chinese characters. *International Journal of Bilingual Education and Bilingualism*, *14*(6), 683–700.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction* (2nd Ed). MIT Press.

Suzuki, Y., & DeKeyser, R. (2017). The interface of explicit and implicit knowledge in a second language: Insights from individual differences in cognitive aptitudes. *Language Learning*, *67*(4), 747–790.

Suzuki, Y., Nakata, T., & Dekeyser, R. (2019). The desirable difficulty framework as a theoretical foundation for optimizing and researching second language practice. *Modern Language Journal*, *103*(3), 713–720.

Swain, M. (2005). The output hypothesis: Theory and research. In *Handbook of research in second language teaching and learning*. Routledge.

Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics*, *16*(3), 371–391.

Taft, M., Castles, A., Davis, C., Lazendic, G., & Nguyen-Hoan, M. (2008). Automatic activation of orthography in spoken word recognition: Pseudohomograph priming. *Journal of Memory and Language*, *58*(2), 366–379.

Tarone, E., & Bigelow, M. (2005). The impact of literacy on oral language processing: Implications for second language acquisition research. *Annual Review of Applied Linguistics*, *25*, 77–97.

Thelwall, R., & Sa'adeddin, A. A. (1990). Illustrations of the IPA: Arabic. *Journal of the International Phonetic Association*, *20*(2), 38–39.

Tushyeh, H. Y. (1996). Linguistic problems facing Arab learners of English. *International Journal of Applied Linguistics*, *111*(1), 109–117.

Tyler, M. D. (2019). PAM-L2 and phonological category acquisition in the foreign language classroom. *A Sound Approach to Language Matters--In Honor of Ocke-Schwen Bohn*, 607–630.

Ullman, M. T. (2005). A cognitive neuroscience perspective on second language acquisition: The declarative/procedural model. *Mind and Context in Adult Second Language Acquisition*.

Ullman, M. T., & Lovelett, J. T. (2018). Implications of the declarative/procedural model for improving second language learning: The role of memory enhancement techniques. *Second Language Research*, *34*(1), 39–65.

Ünal-Logacev, Ö., Żygis, M., & Fuchs, S. (2019). Phonetics and phonology of soft "g" in Turkish. *Journal of the International Phonetic Association*, *49*(2), 183–206.

Upward, C., & Davidson, G. (2011). *The history of English spelling*. Wiley-Blackwell.

Vaid, J. (2022). Biscriptality: a neglected construct in the study of bilingualism. *Journal of*

Cultural Cognitive Science, *6*, 135–149.

Vaid, J., Chen, H.-C., & Rao, C. (2022). Biscriptal bilingualism differentially affects segmentation of cross-language homophones: Evidence from Hindi and English users. *International Journal of Bilingualism*, *26*(1), 13–30.

Vaismoradi, M., & Snelgrove, S. (2019). Theme in qualitative content analysis and thematic analysis. *Forum: Qualitative Social Research*, *20*(3), Article 23.

Van Der Elst, W., Van Boxtel, M. P. J., Van Breukelen, G. J. P., & Jolles, J. (2006). Normative data for the Animal, Profession and Letter M Naming verbal fluency tests for Dutch speaking participants and the effects of age, education, and sex. *Journal of the International Neuropsychological Society: JINS*, *12*(1), 80–89.

van Leussen, J.-W., & Escudero, P. (2015). Learning to perceive and recognize a second language: the L2LP model revised. *Frontiers in Psychology*, *6*, Article 1000.

VanPatten, B. (1996). *Input processing and grammar instruction in second language acquisition*. Ablex publishing.

VanPatten, B. (2004). *Processing instruction: Theory, research, and commentary*. Routledge.

VanPatten, B. (2014). Input processing in adult SLA. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 125–146). Routledge.

VanPatten, B. (2015). Foundations of processing instruction. *IRAL, International Review of Applied Linguistics in Language Teaching*, *53*(2), 91–109.

VanPatten, B., Smith, M., & Benati, A. G. (2020). *Key questions in second language acquisition: An introduction*. Cambridge University Press.

Vaughn, K., Cone, J., & Kornell, N. (2018). A user's guide to collecting data online. *Handbook Res Methods Hum Memory*, 354–373.

Veivo, O., & Jarvikivi, J. (2013). Proficiency modulates early orthographic and phonological processing in L2 spoken word recognition. *Bilingualism: Language and Cognition*, *16*(4), 864–883.

Veivo, O., Järvikivi, J., Porretta, V., & Hyönä, J. (2016). Orthographic activation in L2 spoken word recognition depends on proficiency: Evidence from eye-tracking. *Frontiers in Psychology*, *7*, 1120.

Veivo, O., Porretta, V., Hyönä, J., & Järvikivi, J. (2018). Spoken second language words activate native language orthographic information in late second language learners.

Applied Psycholinguistics, *39*(5), 1011–1032.

Veivo, O., Suomela-Salmi, E., & Järvikivi, J. (2015). Orthographic bias in L3 lexical knowledge. *Language, Interaction and Acquisition*, *6*(2), 270–293.

Venezky, R. L. (1999). *The American way of spelling: The structure and origins of American English orthography*. Guilford Press.

Venezky, R. L. (2005). In search of the perfect orthography. *Written Language and Literacy*, *7*(2), 139–163.

Verhoeven, J., Hirson, A., & Basavaraj, K. (2011). The devoicing of fricatives in southern British English. *International Conference of Phonetic Science XVII*, 2070–2073.

Verhoeven, L., & Perfetti, C. (2017). *Learning to read across languages and writing systems*. Cambridge University Press.

Verhoeven, L., Voeten, M., & Vermeer, A. (2019). Beyond the simple view of early first and second language reading: The impact of lexical quality. *Journal of Neurolinguistics*, *50*, 28–36.

Vouloumanos, A., & Werker, J. F. (2007). Listening to language at birth: evidence for a bias for speech in neonates. *Developmental Science*, *10*(2), 159–164.

Wahba, K. (2021). Models of Arabic pronunciation. In K. Ryding & D. Wilmsen (Eds.), *The Cambridge handbook of Arabic linguistics*. Cambridge University Press.

Watkins, D. (2000). Learning and teaching: A cross-cultural perspective. *School Leadership & Management*, *20*(2), 161–173.

Watson, J. C. E. (2002). *The phonology and morphology of Arabic*. Oxford University Press.

Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language*, *50*(1), 1–25.

Wen, Z. E., Juffs, A., & Winke, P. (2020). Measuring working memory. In P. Winke & T. Brunfaut (Eds.), *The Routledge handbook of second language acquisition and language testing*. Routledge.

Wen, Z., & Li, S. (2019). Working memory in L2 learning and processing. In *The Cambridge handbook of language learning* (pp. 365–389). Cambridge University Press.

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior & Development*, *7*(1), 49–63.

White, L. (2014). Linguistic theory, universal grammar, and second language acquisition. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition* (2nd Ed, pp. 46–65). Routledge.

Winter, B. (2019). *Statistics for linguists: An introduction using R*. Routledge.

Young-Scholten, M. (2002). Orthographic input in L2 phonological development. In P. Bermeister, T. Piske, & A. Rohde (Eds.), *An integrated view of language development: Papers in honor of Henning Wode* (pp. 263–279). Wissenschaftlicher Verlag Trier.

Young-Scholten, M., Akita, M., Cross, N., Robinson, P., & Jungheim, N. O. (1999). Focus on form in phonology: Orthographic exposure as a promoter of epenthesis. *Pragmatics and Pedagogy: Proceedings of the Third PacSLRF*, *2*, 227–233.

Young-Scholten, M., & Langer, M. (2015). The role of orthographic input in second language German: Evidence from naturalistic adult learners' production. *Applied Psycholinguistics*, *36*(01), 93–114.

Zhang, H., & Roberts, L. (2019). The role of phonological awareness and phonetic radical awareness in acquiring Chinese literacy skills in learners of Chinese as a second language. *System*, *81*, 163–178.

Zhang, H., & Roberts, L. (2021). The influence of L1 script directionality and L2 proficiency on Hanzi learning among Arabic and English learners of L2 Chinese. *IRAL, International Review of Applied Linguistics in Language Teaching*, *60*(4), 1039–1062.

Ziegler, J. C., & Ferrand, L. (1998). Orthography shapes the perception of speech: The consistency effect in auditory word recognition. *Psychonomic Bulletin & Review*, *5*(4), 683–689.

Ziegler, J. C., Ferrand, L., & Montant, M. (2004). Visual phonology: the effects of orthographic consistency on different auditory word recognition tasks. *Memory & Cognition*, *32*(5), 732–741.

Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: a psycholinguistic grain size theory. *Psychological Bulletin*, *131*(1), 3–29.

Zwart, F. S., Vissers, C. T. W. M., Kessels, R. P. C., & Maes, J. H. R. (2019). Procedural learning across the lifespan: A systematic review with implications for atypical development. *Journal of Neuropsychology*, *13*(2), 149–182.