

**Developing English Vowel Contrasts: An Analysis of Spanish  
L1 Learners' Speech Production over Time in the UK**

María Gabriela Valenzuela Farías

Doctor of Philosophy

University of York  
Language and Linguistic Science  
December 2022

## **Abstract**

Languages vary in the complexity of their phonemic inventories. Potentially, vowels can possess contrastive features in respect of relative duration, quality, tone and nasality. However, not all languages use all of these features in the production or perception of sounds (Martínez-Celdrán & García, 2019; Ronquest, 2018). For example, Spanish has a simple vowel system in contrast to English which has a more complex vowel system with temporal and spectral differences only. Previous studies have observed that adult Spanish learners struggle to meet the complex target that English presents (Rato & Carlet, 2020). The present research involves the testing of a group of 40 native Spanish-speaking postgraduate students from their arrival at a UK university to the end of their first year of study. In particular, this study examines how the progress in developing the contrasts of English vowel pairs /i:/- /ɪ/; /ɪ/ - /e/ and /u:/ - /ʊ/, is reflected and managed by adaptations towards native-like English vowel pronunciation with respect to vowel quality and temporal features. To this end, Spanish speaking participants were recorded reading a list of words and reading passages at three time points over a year. The analysis was based on formant frequencies and phone duration. The formant values were obtained using Praat software and the degree of separation between the pair of tokens was determined by calculating Euclidean distances. In addition, vocalic changes were judged perceptually by native English lay listeners. Results showed differential rates of progress across members of the group. The differences may be explained by the students' varying exposure to, and engagement with, native English speakers during their year of study in the UK. The results shed light on adult L2 production learning processes and the quantity and quality of time needed for adapting to the new L2 segments.

# Table of Contents

<b>Abstract</b>	<b>2</b>
<b>Table of Contents</b>	<b>3</b>
<b>List of Tables</b>	<b>7</b>
<b>List of Figures</b>	<b>9</b>
<b>Acknowledgments</b>	<b>12</b>
<b>Author's Declaration</b>	<b>13</b>
<b>Chapter 1. Introduction</b>	<b>14</b>
1.1 Aims, research questions and hypotheses	14
<b>Chapter 2. Spanish and English phonological and phonetic features</b>	<b>19</b>
2.1 Introduction	19
2.2 A historical background of the Spanish language	19
2.2.1 Regional varieties of Spanish: Latin American vs Peninsular Spanish	22
2.2.2 Spanish vowel system	24
2.3 British English language: Regional varieties	27
2.3.1 English consonant inventory	27
2.3.2 Standard Southern British English (SSBE) Vowels	28
2.4 Summary	31
<b>Chapter 3. Principal theoretical models for second language production and perception</b>	<b>33</b>
3.1 Introduction	33
3.2.1 The Critical Period Hypothesis (CPH)	34
3.2.2 The Speech Learning Model (SLM)	34
3.2.3 Revised Speech Learning Model (SLM-r)	35
3.2.4 The Second Language Linguistic Perception Model (L2LP)	36
3.2.5 Perceptual Assimilation Model (PAM and PAM-L2)	37
3.3 Summary	38
<b>Chapter 4. Data Collection</b>	<b>41</b>
4.1 Introduction	41
4.2 Participants	41
4.2.1 Language background and nationalities	41
4.2.2 Enrolment status	42
4.2.3 English level	42
4.3 Sample	42
4.4 Recording procedures	43
4.5 Data management	44
4.6 Method	45
4.6.1 Participant recordings	45

4.6.2 Written Language Background Questionnaire -----	48
4.7 Pre-analysis procedures -----	50
4.7.1 Normalisation-----	50
F <sub>n</sub> [v] represents the normalised value of a vowel (i.e., F1/F2). MEAN <sub>n</sub> is the average for the formant (F1/F2) produced by the speaker and S <sub>n</sub> is the standard deviation of the specific formant analysed (Thomas & Kendall, 2007). -----	50
4.7.2 Gender variable -----	51
4.8 Summary -----	51

**Chapter 5. Marking of English Vowel Contrasts over a Year: Overall group performance ----- 52**

5.1 Introduction -----	52
5.1.1 Structure of the chapter -----	52
5.2 Background information -----	52
5.2.1 The acoustics of vowels-----	54
5.2.2 Formants -----	54
5.2.3 Vowel duration -----	54
5.3 The English Vowel Contrasts: /i:/ and /ɪ /; /ɪ / and /e/ and /u:/ and /ʊ/ -----	55
5.4 Analysis -----	58
5.4.1 Formant analysis -----	59
5.4.2 Duration analysis-----	60
5.4.3 Statistical analyses -----	61
5.5 Results -----	62
5.5.1 Formants -----	62
a) Vowel /i:/ -----	62
b) Vowel /ɪ/ -----	63
c) Vowel /u:/-----	63
d) Vowel /ʊ/ -----	64
e) Vowel /e/-----	65
5.5.2 Duration results -----	68
a) Vowels /i:/-/ɪ/-----	69
b) Vowels /u:/-/ʊ/ -----	70
5.6 Summary -----	73
5.6.1 Formant analysis (vowel quality) -----	74
5.6.2 Vowel duration -----	74
5.6.3 General -----	75

**Chapter 6. Marking of English Vowel Contrasts over a Year: Individual performance 76**

6.1. Introduction -----	76
6.1.1 Structure of the chapter -----	76
6.2 Elements affecting L2 vocalic perception and production-----	77
a) Native English speakers -----	77
b) Non-native English speakers-----	78
6.2.1 Exposure, quality and quantity of L2 input-----	80
6.2.2 Input -----	80
6.2.3 Second language experience-----	82
6.2.4 Length of exposure to the target language-----	84
6.2.5 Native language (L1) -----	86

6.3 /u:/and /ʊ/ vowel fronting in British English -----	88
6.4 Euclidean distance -----	90
6.5 Analysis -----	91
6.5.1 Vowel distance analyses (F1-F2) -----	91
6.6 Results -----	91
6.6.1 Vowel separation results -----	92
6.6.2 /i:/ and /ɪ/ vowels results -----	93
6.6.3 F1 and F2 results -----	95
6.6.4 /u:/ and /ʊ/ results -----	97
6.6.5 F1-F2 results -----	99
6.6.6 /ɪ/ and /e/ results -----	101
6.6.7 F1-F2 results -----	103
6.7 /u:/ and /ʊ/ fronting results -----	105
6.7.1 /u:/ - fronting -----	106
6.7.2 /ʊ/ - fronting -----	108
6.8 Vowel duration results -----	109
6.8.1 /i:/ and /ɪ/ vowels -----	110
6.8.2 /u:/ and /ʊ/ vowels -----	113
6.9 Summary -----	116
6.9.1 Formant analysis (vowel quality) -----	116
6.9.2 Vowel duration -----	118
6.9.3 General -----	119

**Chapter 7. Perception by Native British English Speakers ----- 120**

7.1 Introduction -----	120
7.1.1 Structure of the chapter -----	121
7.2 Perception test by Native British English Speakers -----	122
7.2.1 Participants -----	122
7.2.2 Sampling summary -----	122
7.3 Data management -----	122
7.4 Method -----	123
7.4.1 Perception test -----	123
7.4.2 Test design -----	123
7.4.3 Stimuli -----	124
7.4.4 Supporting information -----	125
7.4.5 Testing Procedure -----	128
7.5 Data analysis -----	129
7.6 Results -----	129
7.6.1 /i:/ and /ɪ/ results -----	129
7.6.2 /u:/ and /ʊ/ results -----	130
7.6.3 /ɪ/ and /e/ results -----	131
7.7 Summary -----	133

**Chapter 8. Marking of English Vowel Contrasts over a Year: Individual Performance in Connected Speech ----- 134**

8.1 Introduction -----	134
8.1.1 Structure of the chapter -----	135
8.2 Stimuli -----	135

8.3 Analysis -----	135
8.3.1 Formant analysis -----	136
8.3.2 Vowel distance analysis (F1-F2)-----	138
8.3.3 Vowel duration analysis-----	138
8.4 Results -----	139
8.4.1 Vowel separation results -----	139
8.4.2 /i:/and /ɪ/ results -----	140
8.4.3 F1 and F2 results -----	142
8.4.4 /u:/ and /ʊ/ results -----	144
8.4.5 F1 and F2 results -----	146
8.4.6 /ɪ /and /e/ results-----	149
8.4.7 F1- F2 Results -----	151
8.5 /u:/ and /ʊ/ fronting results-----	154
8.5.1 /u:/-fronting -----	154
8.5.2 /ʊ/- fronting -----	156
8.6 Vowel duration results-----	158
8.6.1 /i:/and /ɪ/ vowels-----	158
8.6.2 /u:/ and /ʊ/ vowels -----	161
8.7 Summary -----	164
8.7.1 Formant analysis (vowel quality) -----	164
8.7.2 Vowel duration -----	166
8.7.3 General -----	167
<b>Chapter 9. Linguistic and social experiences in L2-----</b>	<b>168</b>
9.1 Introduction -----	168
9.1.1 Structure of the chapter -----	168
9.2 Summary of the patterns shown by the participants -----	169
9.2.1 Word List Data – High Group-----	169
9.2.2 Connected Speech Data - High Group -----	170
9.3 Results -----	170
9.3.1 High performing group-----	171
9.3.2 Low performing group -----	174
9.4 Summary -----	178
<b>Chapter 10. Discussion-----</b>	<b>181</b>
10.1 Introduction-----	181
10.1.2 Structure of the chapter -----	181
10.2 Discussion of the findings -----	181
10.3 Limitations and future directions -----	187
10.4 Supplementary information -----	187
<b>References -----</b>	<b>188</b>
<b>Appendices-----</b>	<b>200</b>

## List of Tables

Table 1 Phonemic contrast-----	23
Table 2 Spanish consonant chart-----	24
Table 3 English consonant chart-----	28
Table 4 English vowels in conventional keywords-----	31
Table 5 Main similarities and differences across models-----	39
Table 6 Quality reference values-----	56
Table 7 Duration references values-----	56
Table 8 Summary of the F1 and F2 results-----	67
Table 9 Summary Post-hoc test results-----	72
Table 10 Vowel duration values-----	75
Table 11 Comparative table of F2 frequencies from different studies-----	90
Table 12 Summary T1-T3-----	94
Table 13 Overall results showing the number of speakers producing higher or lower F1 values from Time 2-3 -----	96
Table 14 Overall results showing the number of speakers producing higher or lower F2 values from Time 2-3 F2 -----	96
Table 15 Summary T1-T3-----	98
Table 16 Overall results showing the number of speakers producing higher or lower F1 values from Time 2-3 -----	100
Table 17 Overall results showing the number of speakers producing higher or lower F2 values from Time 2-3 -----	101
Table 18 Summary T1-T3-----	102
Table 19 Overall results showing the number of speakers producing higher or lower F1 values from Time 2-3 -----	104
Table 20 Overall results showing the number of speakers producing higher or lower F2 values from Time 2-3 -----	104
Table 21 Summary T1-----	110
Table 22 Summary T2-----	111
Table 23 Summary T3-----	112
Table 24 Summary T1-----	113
Table 25 Summary T2-----	114

Table 26 Summary T3-----	115
Table 27 Confusion matrix for vowels. The row represents the target vowels. The column represents the listeners' responses -----	133
Table 28 Total number of tokens per vowel across a series of time points-----	137
Table 29 Summary T1-T3-----	141
Table 30 Overall results showing the number of speakers producing higher or lower F1 values from Time 2-3 -----	143
Table 31 Overall results showing the number of speakers producing higher or lower F2 values from Time 2-3 -----	143
Table 32 Summary T1-T3-----	145
Table 33 General results showing the number of speakers producing higher or lower F1 values from Time 2-3 -----	147
Table 34 General results showing the number of speakers producing higher or lower F2 values from Time 2-3 -----	147
Table 35 Summary T1-T3-----	150
Table 36 Overall results showing the number of speakers producing higher or lower F1 values from Time 2-3 -----	152
Table 37 Overall results showing the number of speakers producing higher or lower F2 values from Time 2-3-----	152
Table 38 Summary T1-----	158
Table 39 Summary T2-----	159
Table 40 Summary T3-----	160
Table 41 Summary T1-----	161
Table 42 Summary T2-----	162
Table 43 Summary T3-----	163
Table 44 Word list-----	169
Table 45 Connected speech -----	170
Table 46 Self- reported most difficult skill-----	171
Table 47 Time spent at the university in academic activities-----	173
Table 48 Self- reported most difficult skill-----	175
Table 49 Time spent at the university in academic activities-----	176
Table 50 Summary Table-----	179



## List of Figures

Figure 1 Spanish language as an official language around the world-----	22
Figure 2 Spanish vowel classification-----	25
Figure 3 Articulation of Spanish vowels-----	26
Figure 4 Contemporary RP-----	29
Figure 5 Contemporary RP and Spanish vowels-----	30
Figure 6 Recording equipment-----	44
Figure 7 Sample of a recording session-----	48
Figure 8 Sample of the analysed word. The word uttered in this extract is the word ‘best’. The menu on the right was used to log the formants, vowel and source for retrieval and analysis -----	60
Figure 9 F1-F2 means for /i:/-----	62
Figure 10 F1-F2 means for /ɪ/-----	63
Figure 11 F1-F2 means for /u:/-----	64
Figure 12 F1-F2 means for /ʊ/-----	65
Figure 13 F1-F2 means for /e/-----	66
Figure 14 Mean duration of vowel /i:/ for Time 1, 2, and 3-----	69
Figure 15 Mean duration of vowel /ɪ/ for Time 1, 2, and 3-----	69
Figure 16 Vowel contrast duration /i:/-/ɪ/ for Time 1, 2, and 3-----	70
Figure 17 Mean duration of vowel /u:/ for Time 1, 2, and 3-----	71
Figure 18 Mean duration of vowel /ʊ/ for Time 1, 2, and 3-----	71
Figure 19 Vowel contrast duration /u:/-/ʊ/ for Time 1, 2, and 3-----	72
Figure 20 Possible assimilation patterns for British English produced by Spanish speakers -----	87
Figure 21 Vowel distance variations between /i:/ and /ɪ/ from Time 1-3 -----	94
Figure 22 Summary in percentages about the proportion of speakers who fell into either of the two categories -----	95
Figure 23 Vowel distance variations between /u:/ and /ʊ/ from Time 1-3 -----	98
Figure 24 Summary in percentages about the proportion of speakers who fell into either of the two categories -----	99
Figure 25 Vowel distance variations /ɪ/ and /e/ from Time 1-3 -----	102

Figure 26 Summary in percentages about the proportion of speakers who fell into either of the two categories -----	103
Figure 27 Changes in the F2 values for /u:/ between T1 and T3 -----	107
Figure 28 Summary in percentages about the proportion of speakers who fell into either of the three categories -----	107
Figure 29 Changes in the F2 values for /ʊ/ between T1 and T3 -----	108
Figure 30 Summary in percentages about the proportion of speakers who fell into either of the three categories -----	109
Figure 31 Duration difference between /i:/ and /ɪ/ at Time 1 (voiceless final consonants)-	110
Figure 32 Duration difference between /i:/ and /ɪ/ at Time 2 (Voiced final consonants)----	111
Figure 33 Duration difference between /i:/ and /ɪ/ at Time 3 -----	112
Figure 34 Duration difference between /u:/ and /ʊ/ at Time 1-----	113
Figure 35 Duration difference between /u:/ and /ʊ/ at Time 2 -----	114
Figure 36 Duration difference between /u:/ and /ʊ/ at Time 3 -----	115
Figure 37 Picture that shows the design process of the perception test -----	124
Figure 38 Sample of the final format of the online test -----	126
Figure 39 Sample of the final online format of for the background information questions-	127
Figure 40 Overall correct identification for /i:/-/ɪ/ by N.E speakers-----	130
Figure 41 Overall correct identification for /u:/-/ʊ/ by N.E. speakers-----	131
Figure 42 Overall correct identification for /e/ by N.E speakers-----	132
Figure 43 Formant analysis sample. The speech uttered in this extract is the word ‘attempt’. The menu on the left is used to log the formants, vowel category and source - -	138
Figure 44 Vowel distance variations between /i:/ and /ɪ/ from Time 1-3 -----	141
Figure 45 Summary in percentages about the proportion of speakers who fell into either of the two categories -----	142
Figure 46 Vowel distance variations between /u:/ and /ʊ/ from Time 1-3 -----	145
Figure 47 Summary in percentages about the proportion of speakers who fell into either of the two categories -----	146
Figure 48 Vowel distance variations between /ɪ /and /e/ from Time 1-3 -----	150
Figure 49 Summary in percentages about the proportion of speakers who fell into either of the two categories -----	151
Figure 50 Changes in the F2 values for /u:/ between T1 and T3-----	155

Figure 51 Summary in percentages about the proportion of speakers who fell into either of the three categories -----	155
Figure 52 Changes in the F2 values for /ʊ/ between T1 and T3 -----	156
Figure 53 Summary in percentages about the proportion of speakers who fell into either of the three categories -----	157
Figure 54 Duration difference between /i:/ and /ɪ/ at Time 1 -----	158
Figure 55 Duration difference between /i:/ and /ɪ/ at Time 2 -----	159
Figure 56 Duration difference between /i:/ and /ɪ/ at Time 3 -----	160
Figure 57 Duration difference between /u:/ and /ʊ/ at Time 1 -----	161
Figure 58 Duration difference between /u:/ and /ʊ/ at Time 2 -----	162
Figure 59 Duration difference between /u:/ and /ʊ/ at Time 3 -----	163

## Acknowledgments

I would like to thank my supervisor, Peter French, for his constant support, encouragement, patience (a lot), friendship, sense of humour and knowledge throughout this thesis process and during the COVID-19 pandemic.

Thanks to my TAP member Carmen Llamas for her advice; to Dominic Watt for his detailed feedback and to Phillip Harrison and Huw Jones for their help.

I would like to acknowledge my friend Sascha Schäfer for his great disposition to help me, and Hanna Klein for proofreading the chapters of this thesis.

Thanks to my ‘lockdown-bubble’(Andrea and Paulina) for accompanying and ‘*adopting*’ me during the hard times of lockdown. Thanks to all my good friends in York and to my family in Chile, especially to my parents, brother (J. Andrés), my sister in-law (Karen) and my lovely nieces, Isidora and Camila for their constant support.

A special acknowledgement to my Tata, John Carter and tío Abraham.

My gratitude to the Chilean government for funding this research through ANID.

## **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 an award at this, or any other, University. All sources are acknowledged as References.

## Chapter 1. Introduction

Languages differ in their complexity of phonemic inventories. An important task for linguistics is to therefore analyse the range of phonemic variation across different languages and to determine what this range may mean for foreign learners. English is today's *lingua franca* and non-native English learners across all ages and with different first languages (L1s) are faced with some difficulties in cross-language speech production.

The acquisition of the phonological system of English as a second language (L2) has been the subject of different experimental studies that debate the degree to which phonemic and phonetic learning in adults remains feasible through the L2 experience. Different factors that affect L2 phoneme learning have been described in theories and language learning models. For example, the relationship between L1 and L2 phoneme systems (Best & Tyler 2007; Flege, 1995), the age of L2 learning, the length of L2 exposure and L1 use (Flege, 1999; Flege et al., 2021) have all been proposed as significant factors. Despite proposals that a range of influences can facilitate or hinder the learning of L2 segments (especially in adults), nearly all authors agree that (even though age matters) adults are not excluded from developing phonological competence in an L2 (Munro & Derwing, 2008).

Speakers from different L1 backgrounds may face different problems when learning English vowel phonemes. A major factor here concerns the degree of difference between the L1 vowel inventory and that of English. For example, Spanish speakers have an L1 with a small vowel inventory (5 phonemes) and this tends to present them with problems in mastering contrasts found in the more extended English vowel system. This research considers a group of Spanish participants' progress, both individually and *qua* group, in the production of a set of L2 English vowel contrasts over the course of a year in an English university.

### 1.1 Aims, research questions and hypotheses

The purpose of the thesis is to longitudinally examine the progress of adult Spanish-speakers in distinguishing between pairs of English vowels /i:/ and /ɪ/, /ɪ/ and /e/ and /u:/ and /ʊ/, in terms of both quality and duration. The external factors that might influence speakers' development of the pair contrasts over the course of a year are considered from information the participants provided in response to a questionnaire.

The specific phonetic target pronunciations for the participants are difficult to define owing to the wide – but unknown – range of English accents that they were differentially exposed to. However, the difficulties in this respect are to an extent mitigated by the fact that three of the five vowels selected for the study are phonologically ‘short’ vowels of British English and therefore subject to least regional and social variation (French et al., 2008; Wells, 1982).

The following are the five research questions the research seeks to address and the associated hypotheses:

Q1. Does the group as a whole show a movement towards producing a phonemic contrast between /i:/ and /ɪ/, /ɪ/ and /e/ and /u:/ and /ʊ/ over the course of their year in England?

H1. Given the immersion of the group in a range of English social, commercial and educational settings, one would expect to find adaptation to L1 patterns, as manifest in these vowel contrasts and pronunciations.

Q2. If so, how do the learners mark the developing contrast between members of the front vowel pairs and of the back vowel pair, by vowel quality as indexed by the first and second formant values, or by length, or by both?

H2. Given the findings of some previous studies of Spanish learners, which have found movement towards native English vowel quality over shorter periods of time than a year (see Flege, 1988; Munro & Derwing, 2008), one would expect to find - minimally - changes in quality; there is no hypothesis as to whether the present participants would also begin to mark the contrasts durationally.

Q3. Over time, do some individual speakers develop more marked contrasts than do others; and, if so, what are the specific experiential L2 exposure/engagement factors associated with them?

H3. Given that learners will have different degrees of exposure to, and experience of, English speaking contexts - educational, commercial and social - it is anticipated that

there will be differential rates of progression. While it is possible to hypothesise at a general level that the overall amount of exposure to native English, and that active engagement with native English speakers rather than simply passive exposure to the language, might be facilitative of progression. At this point there is no hypothesis as to which specific contexts will be most involved.

Q4. Are the contrasts - as determined by acoustic measurements - perceptible by native English speakers?

H4. It is hypothesised that the contrasts will be perceived by L1 native speakers, and, based on studies by, for example, Flege et al. (1997) and Hillenbrand et al. (2000), those that are marked by vowel quality will be more perceptually salient for L1 English speakers than those marked just by vowel length differences.

Q5. Do the patterns found in the present data have a bearing on extant conceptual models of L2 phonological/phonetic learning?

H5. Given the fact that the different models share similarities and differences regarding L2 development, it is hypothesised that the patterns found in the present data may contribute to the conceptual models that emphasise adults' L2 production over children's L2 system development, and to models where learning English vowel contrasts is specifically addressed, as in the Speech Learning Model of Flege (1995) and the Second Language Linguistic Perception Model of Escudero (2005), respectively.

In addressing these questions, it is anticipated that the research might provide useful insights into L2 phonemic development. Further, the findings support the view that native English speakers rely on the quality of the vowels as a primary source of cues to categorical phoneme perception, and that duration plays a secondary role. Finally, the results shed new light on previous ideas about the phonetic/phonological learning process, which proposes that adult L2 speakers progress fast initially and then plateau (see Flege et al., 1992).

This thesis does not begin with a wide-ranging literature review. The existing studies relevant to it overall span a number of different areas. For convenience to the reader and for textual



coherence, therefore, literature relevant to each chapter is referred to and explained within the chapter itself.

The structure and organisation of the thesis is as follows:

**Chapter 2** presents background information on the historical spread and current global distribution of the Spanish language. The main features of Spanish and English phonetics and phonology are presented and compared, with an emphasis on the differences between their vowel inventories.

**Chapter 3** sets out the major theoretical models of L2 learning in order of their chronological emergence and development. It considers the factors that promote and inhibit efficacious learning and brings the constituent propositions to bear on the research.

**Chapter 4** describes the fieldwork techniques used to obtain the speech data on which the thesis is based. The participants, the recording procedures and the stimuli used in the different recording sessions are explained. In addition, a description of the language background questionnaire (implemented after each of the recording sessions) is provided.

**Chapter 5** contains the results of the group performance over time (a general view) in the production of the English vowel pairs contrast /i:/ and /ɪ/, /ɪ/ and /e/, /u:/ and /ʊ/, with respect to both spectral and temporal features. The statistical analysis presented in this chapter was conducted using citation data gathered from word list reading at three time points.

**Chapter 6** reports on the performance of high- and low-performing individuals in the group, again from the word list data and again with respect to spectral and durational analyses of the vowels under study. The identification of high and low performers is based on the degree of separation participants made between members of the English vowel pairs across the three time points.

**Chapter 7** presents a perception test carried out with native British English raters. This additional tool was adopted to determine empirically whether the distinctions, established analytically, made by the participants had consequences for how the productions are heard by native speakers.

**Chapter 8**, like Chapter 6, analyses the results obtained from each individual speaker, but in respect of running speech – reading of phonetically-balanced passages – rather than isolated item-by-item word list data. The results of the word list and connected passage elicitations are compared.

**Chapter 9** is concerned with explaining the possible reasons for the differences between high- and low-performing individuals. In so doing, information from the language background questionnaire is drawn upon. Given that the participants had differential exposure to native spoken English in different communicative contexts over the course of the year, the question is addressed of which types of exposure and which contexts were most conducive to increases in performance.

**Chapter 10** presents a general discussion of the findings overall. The limitations and directions for future research are set out.

## **Chapter 2. Spanish and English phonological and phonetic features**

### **2.1 Introduction**

An important aim of linguistics is to comprehend and document the range of variation across languages. At the phonological level, it is recognised that languages vary in terms of the complexity of the phoneme inventory (vowels and consonants), with Spanish and English being, respectively, examples of languages with simple and more complex phonemic systems, including their vowel inventories. Potentially, vowel phonemes can possess contrastive features in respect of relative duration, quality, tone and nasality; however, not all languages use all of these features in the production or perception of contrastive segments (Martínez-Celdrán & Elvira-García, 2019; Ronquest, 2018). For example, Spanish is a language with a small and ‘uncomplicated system’, as it has only 5 monophthong vowel phonemes, which are distinguished from one another by quality alone. Standard Southern British English has 12 different monophthongs involving both quality and durational contrasts. In terms of consonants, Spanish has 19 consonant phonemes: 6 plosives, 2 affricates, 3 nasals, 1 tap, 1 trill, 4 fricatives and 2 laterals (Martínez-Celdrán et al., 2003), whereas English has 24 consonant phonemes: 6 plosives, 2 affricates, 3 nasals, 9 fricatives, 3 approximants, and 1 lateral (Upton et al., 2001).

The above cross-linguistic variances are presented in this chapter, which starts with a brief historical review of the spread and distribution of the Spanish language, followed by a description of the articulatory and acoustic properties of Spanish consonants and vowels (including regional varieties). The focus then shifts to the Standard Southern British English (near) counterparts of those phonemes and ends with a summary of the two.

### **2.2 A historical background of the Spanish language**

The language of the Spanish Peninsula has been modelled by different historical events including the Roman invasion, the Muslim Moorish conquest, the re-conquest of the Iberian Peninsula and the invasion/discovery of America *inter alia*. Spanish originally evolved from Vulgar Latin. Later, the standard language followed the rules of the Castilian variety spoken

in the Kingdom of Castile (Pharies, 2007). Castellano or Castilian was made the official language of Spain under the rule of the Catholic monarchs (Queen Isabella of Castile and King Ferdinand II of Aragón). It was promoted and codified under the *Gramática de la Lengua Castellana* (grammar of the Castilian language) and eventually became the language of the elite and educated people (Cano, 1995). Castilian was further spread and used in the Spanish colonies (Amstutz, 2020) during the colonisation of the Americas. The Spanish spoken in the colonies remained close to the standard Spanish spoken in the Peninsula. During the 18th century, to ensure the dominance of Castilian, the use of Amerindian languages in administrative and educational institutions was prohibited (Amorós-Negre, 2016). As a consequence (of Spanish colonisation) 18 countries on the American continent adopted Spanish as their official language (Mexico, Nicaragua, Panama, Paraguay, Colombia, Costa Rica, Cuba, Dominican Republic, Puerto Rico, Ecuador, El Salvador, Guatemala, Honduras, Venezuela, Uruguay, Argentina, Bolivia, Perú, Chile) and the African state of Equatorial Guinea, and the Philippines.

The creole elite (in the new republics) owned economic and linguistic capital. They were monolingual Spanish speakers, with a strong connection to Spain, who considered indigenous languages as an impediment to economic growth (Amorós-Negre, 2016). Education in the new republics was given in Spanish and provided by the Catholic church; the development of bilingualism (Spanish-indigenous languages) was therefore slow, and Spanish maintained its predominance as the official language with both linguistic and economic capital (Melià, 1992). During colonial times, and following the independence of different countries, the standard Spanish spoken in the Peninsula was considered an ‘exemplar’ model which was taught in schools and spoken by educated people and the creole elite (Lara, 2011). In 1773, in order to maintain the hegemony of the language, the Marquis of Villena founded the *Real Academia Española* (RAE - Royal Spanish Academy), an institution used as a linguistic reference point to this day. The main objective of the RAE is to guarantee that changes in the language, due to the constant adaptations to the needs of its speakers, do not break the essential linguistic unity in the Hispanic sphere. The RAE is responsible for establishing and disseminating criteria for ‘the correct’ use of the language in the Spanish-speaking communities. In this way, the consolidated bases achieved over centuries can be maintained (Real Academia Española, 2019).

Following and before the dictatorship of Franco (1939-1975), several dialects were spoken in the Iberian Peninsula including Aragonese, Asturian, Basque, Caló, Catalán, Valencian and Galician, among others. However, during Franco's regime minority languages and associated cultural practices (such as traditional dance) were prohibited in schools, government settings, and public events. Indeed, speaking any language other than Spanish was punishable by imprisonment and fines (Vann, 1999). Castilian was proclaimed (again) as the official language of Spain in order to promote the unity of the country (Sala & Posner, 2021). Following Franco's administration, Castilian continued to be the official language of Spain. However, the existence of co-official languages in some parts of Spain (such as in Catalunya, Aragón, and País Vasco, amongst other regions) has been constitutionally recognised and accepted within the country (Taylor, 2022). Also, in Latin America some indigenous languages, such as Guaraní (Paraguay) and Quechua (Ecuador) have been proclaimed as national languages in addition to Spanish, but these languages have not gained the same status as Spanish (Amorós-Negre, 2016).

It is possible that the efforts to preserve the Spanish language (with a relative low degree of variation throughout its history) has had impressively positive results in Spanish-speaking communities. In 2004, the RAE and the Association of the Spanish Language Academies (ASALE, which includes Latin American countries, the Philippines, Equatorial Guinea and Spain) created the New Pan-Hispanic Language Policy to reinforce the sense of unity in Spanish-speaking countries. This policy incorporated both the lexical and grammatical changes that Spanish has undergone outside the Peninsula with the ultimate goal of obtaining a single linguistic model that integrates the polycentric varieties of the language (Amorós-Negre, 2016).

Nowadays, Spanish is spoken as a native language by more than 400 million people worldwide, with Mexico being the country with the most speakers (in total 120 million+) followed by Colombia (48 million+speakers) and Spain (42 million+speakers) (Instituto Cervantes, 2017).

The following figure illustrates where Spanish is currently spoken as an official language.

**Figure 1**

*Spanish language as an official language around the world*



Source: Encyclopædia Britannica, Inc. 2021. (<https://www.britannica.com/topic/Spanish-language#/media/1/558113/239454>)

### **2.2.1 Regional varieties of Spanish: Latin American vs Peninsular Spanish**

Given the institutional measures described above taken to secure homogeneity and standardisation, Spanish has only evolved to a very limited extent in different ways across time in the regions where it is spoken. According to Lipski (2012), there have been many factors giving rise to these minor dialectal variations of Spanish, such as education, migrations, missionary activities in rural areas, language policies within countries, and language contact.

Broadly speaking, the number of Spanish phonemes (vowels and consonants) remains equal across all dialects, although there are some minor phonetic and phonological variations (Lipski, 2012). For example, (Latin) American Spanish generally, in contrast to the Peninsular

Spanish, has 18 consonant phonemes, lacking the voiceless interdental fricative /θ/, and in some cases preserving the palatal lateral /ʎ/, a phoneme which is disappearing in Spain (Lipski, 2012; Salcedo, 2010). In Peninsular Spain, the /θ/ and /s/ are two separate phonemes; however, the opposition of /θ/-/s/ does not occur in Andalucía, the Canary Islands and Latin America. Thus, in those varieties, words such as *losa* /'losa/ 'carved stone' or *loza* /'loθa/ 'crocery', *casa* /'kasa/ 'house' or *caza* /'kaθa/ 'hunting' are not minimal pairs, all having only /s/. In addition, the palatal lateral /ʎ/ is now rarely produced by Peninsular Spanish speakers, giving way to /j/. However, this distinction is generally maintained in Latin American countries such as Paraguay, Ecuador, Perú, Bolivia, and Colombia where words such as *cayó* /ka'jo/ 'fell down' and *calló* /ka'ʎo/ 'went silent' are minimal pairs (Hualde, 2005; Lipski, 2012). The following table provides more information on the geographical distribution of the above two consonantal features.

**Table 1.**

*Phonemic contrast*

1. /s/ vs /θ/ Only in Northern-Central Peninsular Spanish (northern and central Spain)		
/θ/	z, c(e,i)	<i>cena</i> /θéna/, <i>escena</i> /esθéna/, <i>azul</i> /aθúl/
/s/	s	<i>saco</i> /sáko/ 'bag'
2. /j/ vs. /ʎ/ Only in parts of Spain, the Andean region and Paraguay		
/j/	y	<i>vaya</i> /bája/ 'that s/he/I go'
/ʎ/	ll	<i>valla</i> /báʎa/ 'fence'

(From Hualde, J. 2005, p.8)

The main 19 consonant phonemes are shown in Table 2.

**Table 2***Spanish consonant chart*

	Bilabial	Labiodental	Dental	Alveolar	Palatal	Velar
Plosive	p   b		t   d			k   g
Affricate					tʃ   ʝ	
Nasal	m			n	ɲ	
Tap or flap				r		
Trill				r		
Fricative		f	θ	s		x
Lateral approximant				l	ʎ	

Note: left-side sounds are voiceless, right-side sounds are voiced. (From Martínez-Celdrán et al., 2003).

Although Spanish segments present minimal variation among dialects, intonation is a clearly distinguishable suprasegmental feature which differs across Spanish varieties. The differences are generally explained due to the language contact with indigenous languages, European settlements in the Americas, and the development of bilingualism (i.e., Spanish - native indigenous language) in some countries (see Colantoni & Gurlekian, 2004; Hualde & Prieto, 2015; O'Rourke, 2005).

### 2.2.2 Spanish vowel system

The Spanish vowel system has been described as a simple and symmetrical one because it possesses only five tense, short contrasting monophthongs: /i/, /e/, /a/, /o/, /u/ - with minimal variation between dialects<sup>1</sup> (Salcedo, 2010). According to Hualde (2005), these vowels may be classed in accordance with three main dimensions: tongue height (high, mid, low), tongue backness (front, central, back), and lip rounding (rounded and unrounded).<sup>2</sup> They can occur in stressed and in unstressed positions without being reduced or dropped (Martínez- Celdrán & Elvira-García, 2019). Close front [i] can be found in words such as *misa* 'mass', the close-mid-front [e] in words such as *mesa* 'table', the opencentral [a] is produced in *masa* 'dough',

<sup>1</sup> Some dialects in America show devoicing of vowels in unstressed syllables especially when vowels are after dental sounds such as 'pues' [ps](Martínez- Celdrán & Elvira-García, 2019), but these are realisational (allophonic) rather than systemic variations.

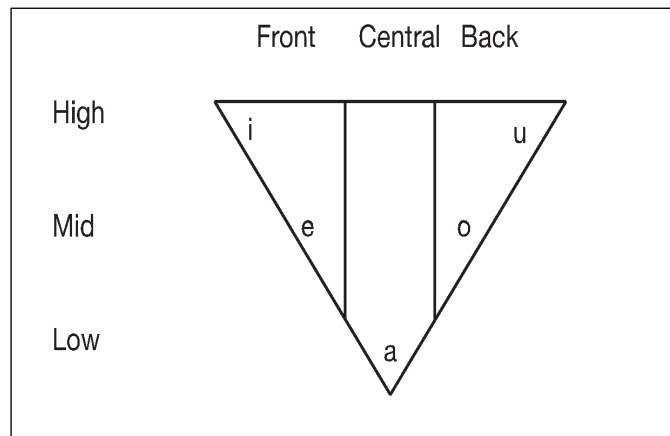
<sup>2</sup> Unrounded vowels are the front and central ones, while back vowels are considered rounded.



the close back [u] as in *musa* ‘muse’, and the close-mid-back [o] in words as in *moza* ‘girl’. The following figure shows the triangular articulatory arrangement for Spanish vowels within vowel space.

**Figure 2**

*Spanish vowel classification*

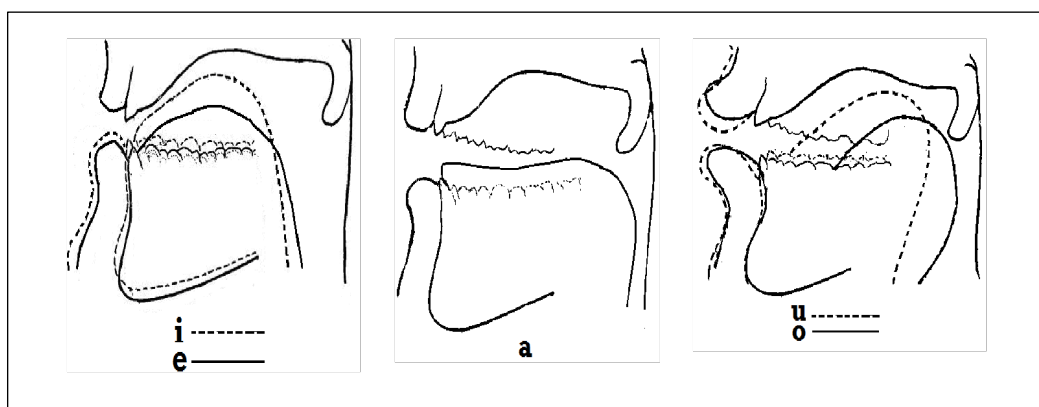


(From Salcedo, 2010, p.199)

In terms of production, no contact with articulators is made; lip rounding, duration and nasalization are not used to distinguish contrastive segments. In brief, front vowels are made with unrounded lips and positioning the tongue in the front area of the oral cavity; the central vowel is produced with unrounded lips and low steady tongue, whereas back vowels are articulated with rounded lips and the tongue dorsum in open approximation with the velum (Matínez- Celdrán & Elvira-García, 2019). The following figure shows the articulation of Spanish vowels.

### Figure 3

#### *Articulation of Spanish vowels*



(From Martínez- Celdrán & Elvira-García, 2019)

The different tongue positions for the vowels determines their auditory quality; this can also be estimated acoustically from their constituent areas of high energy within the frequency spectrum, i.e., their formants. Ladefoged (2006) states three main formants (F1, F2, F3) for vowel distinction, with the first two (F1, F2) being the ones used to indicate the position of vowels within the acoustic space. Concisely, F1 frequency is inversely related to tongue height (the higher the vowel, the lower the F1 value) and F2 is inversely related to tongue backness (the more back the vowel, the lower the F2 value).

With respect to duration, Spanish vowels possess intrinsic duration; that is, in common with other languages, high vowels (e.g. /i/, /u/) are mostly produced shorter than low vowels (Chládková et al., 2011). Vowel lengthening is not used to produce or perceive vocalic contrast (as may be the case in English); however, variations of duration do exist, and, in addition to their being related to vowel height, they are mainly affected by consonant context, speaker gender, pragmatic purpose, and syllable stress (Matínez- Celdrán & Elvira-García, 2019; Ronquest, 2018).

In summary, it has been shown that overall, the Spanish vowel system shows consistency and minimal variation across regions, the quality of the vowels is relatively steady during articulation processes, and quantity is not used to signal contrast between segments. The symmetrical 5-monophthong system is the commonest cross-linguistically.

To explore this contrast in phonemic inventories between Spanish and English, the following section covers a brief description of British English. The first part covers regional varieties of British English and continues with an overview of the Standard Southern British English vowel system.

### **2.3 British English language: Regional varieties**

English is spoken globally – whether as a first, second, or foreign language. Its wide use in different areas where it is not an official language means it increasingly serves as a *lingua franca*. Thus, many native and non-native varieties can be observed. Within the language there are different dialects; varieties very often differ in the vowels that speakers produce, and this has traditionally been associated with geography, social class, gender and age. Overview studies such as the volumes of *Accents of English* by Wells (see Wells, 1982) have provided historical and phonological accounts of social and regional accent variation over time in the UK and overseas. Other publications have documented geographic and social changes in pronunciation, grammar and vocabulary of more than 30 varieties of English in the UK and Ireland (see Hughes et al., 2012). And numerous sociolinguistic studies have confirmed that age, gender, social class and geography are factors that have a great impact on how dialects, in this case of British English, may differ (see Haddican et al., 2013, Llamas, 2007). However, many English speakers are able to adapt their speech for specific circumstances or purposes, resulting in their being considered bidialectal (Ladefoged & Disner 2012).

#### **2.3.1 English consonant inventory**

Generally speaking, vowels show greater variation across varieties of English than do consonants. It is therefore easier to make generalisations about English consonants (Ladefoged, 2006; Ladefoged & Ferrari, 2012). English consonants may be classified under the same three parameters as the Spanish ones: place of articulation, manner of articulation, and voicing (Roach, 2004). The following table presents the 24 English consonant phonemes.

**Table 3***English consonant chart*

	Bilabial	Labiodental	Dental	Alveolar	Post-alveolar	Palatal	Velar	Glottal
Plosive	p b			t d			k g	
Affricate					tʃ dʒ			
Nasal	m			n			ŋ	
Fricative		f v	θ ð	s z	ʃ ʒ			h
Approximant	(w)				r	j	w	
Lateral approximant				l				

Note: where consonants appear in pairs, the left-hand symbol is voiceless, the right-hand one voiced (from Roach, 2004).

### 2.3.2 Standard Southern British English (SSBE) Vowels

The question of how many vowels are in present English is difficult to answer with precision. The vowel inventory varies somewhat depending on the dialect of English being described. For example, General American English (GAE) has 14 -15 vowel phonemes, while SSBE is considered to have around 20 (including diphthongs) (Ladefoged & Disner, 2012). For the purpose of this research, and to narrow down the description of English vowels, the focus is placed on SSBE. SSBE is now used to refer to the accent that tends to accompany to 'Standard British English'; it is considered the present-day equivalent of what has been referred to as Received Pronunciation (RP), or BBC English; it is mainly associated with the upper and middle classes from the south of England.

SSBE is not exactly the same as RP. It is not my intention to enter into debate regarding the RP recession or status; instead, I limit myself to use the term SSBE to refer to the accent of England which is currently considered to be the standard. However, I adopt the term RP, contemporary RP, or SSBE when discussing literature in which the author herself/himself uses it.

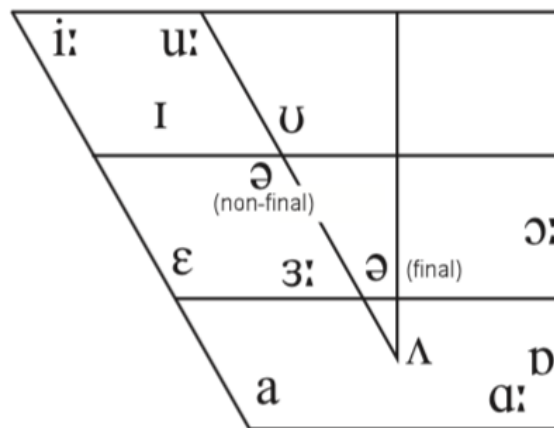
SSBE features 12 monophthongs: five long ones /i:/, /ɜ:/, /ɑ:/, /ɔ:/, /u:/ and seven short vowels /ɪ/, /e/, /æ/, /ʌ/, /ɒ/, /ʊ/, /ə/. The terms ‘long’ and ‘short’ are not to be considered as referring to a fixed duration, because phonetically the length of a vowel depends on the phonological context and the syllable stress (Roach, 2004). For example, while /ɪ/ is always short, /i:/ is not always long. Specifically, the following rule applies:

- i. /i:/ → [i:] open syllable
- ii. /i:/ → [i·] syllable closed by voiced consonant
- iii. /i:/ → [i] syllable closed by voiceless consonant (see Roach, 2009)

The position within the vowel space of contemporary RP vowels is shown in the following Figure:

**Figure 4**

*Contemporary RP*

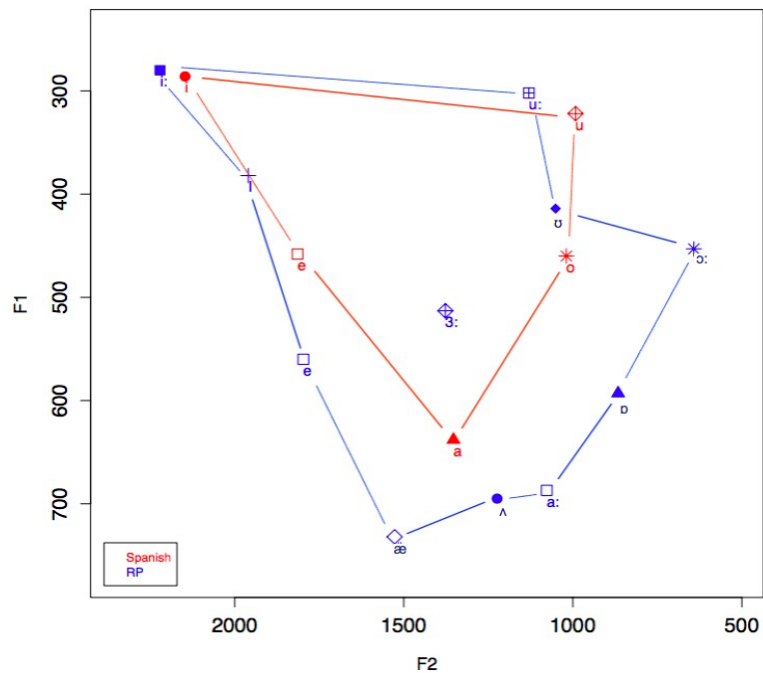


(Source: *English accents and dialects*, Hughes, et al, 2012, 5<sup>th</sup> ed., p.53)

In order to illustrate the complex distribution of the British English vowels in F1-F2 space, the contemporary RP values have been compared to those of the Spanish vowels (Figure 5).

**Figure 5**

*Contemporary RP and Spanish vowels*



Note: The contemporary RP values were taken from Deterding (1990 as cited in Deterding 1997) and the Spanish values from Bradlow (1994).

In recent years many - mainly British - authors have adopted the convention, following Wells (1982), of not using phonetic symbols to represent English vowel phonemes, but 'headwords' for the lexical sets that contain them. Although I have chosen to retain the convention of phonetic symbolisation, the table below shows the relationship of the phonetic symbol convention against the headword option for the vowels of RP/SSBE.

**Table 4***English vowels in conventional keywords*

<b>Keyword</b>	<b>Vowel</b>	<b>Keyword</b>	<b>Diphthong</b>
KIT	/ɪ/	FACE	/eɪ/
DRESS	/e/	GOAT	/əʊ/
TRAP	/æ/	PRICE	/aɪ/
LOT	/ɒ/	CHOICE	/ɔɪ/
STRUT	/ʌ/	MOUTH	/aʊ/
FOOT	/ʊ/	NEAR	/ɪə/
BATH	/ɑː/	SQUARE	/eə/
CLOTH	/ɒ/	CURE	/ʊə/
NURSE	/ɜː/		
FLEECE	/iː/		
PALM	/ɑː/		
THOUGHT	/ɔː/		
GOOSE	/uː/		
START	/ɑː/		
NORTH	/ɔː/		
FORCE	/ɔː/		
HAPPY	/ɪ/		
lettER	/ə/		
commA	ə/		

(Adapted from Ogden, 2009, p.67)

## 2.4 Summary

This chapter has described the historical background of the spread and development of Spanish across the world. Spanish retains a remarkable degree of homogeneity despite its wide geographical spread and the amount of time it has been present in geographically non-contiguous areas. In particular, the Spanish vowel system is fairly constant internationally. This lack of vowel variation has been a great asset for the present research, since it has made it possible to include speakers from various countries in the knowledge that they will all approach English from more or less the same starting point.

Spanish and English differ in terms of their vowel inventories, with English having a complex system and Spanish a simple system. Also, as indicated in Figure 5, while there are approximations between some of the Spanish and English vowel phonemes, the normative realisations in terms of vowel quality may not entirely converge. The importance of these differences to the present research is that vowel contrasts that exist in English but not in Spanish may present Spanish learners of English with particular problems, as may adjustments in pronunciation from Spanish to English norms in respect of phonemes that stand in near correspondence across the two languages. This view is developed and explained

further in Chapters 5 and 6. More immediately, however, I present some major, general theoretical models of second language acquisition that may throw light on the issue and the findings (Chapter 3), and an exposition of the data elicitation and recording methods (Chapter 4).



## **Chapter 3. Principal theoretical models for second language production and perception**

### **3.1 Introduction**

A great deal of research has been undertaken on how learners of English as a second language (L2) acquire English vowel contrasts. In particular, studies of the differences between children (before puberty) and adults in learning an L2 have been undertaken (Hu, 2016). Flege (1999) explains that the main difference between adults and children is that the latter are more capable of forming new L2 categories because their first language (L1) is not fully developed; that is, the phonological categories for L1 vowels and consonants which are starting to be formed and defined during childhood lay the foundation to process phonetically different vowels and consonants in an L2, blocking the perceptual assimilation of L1-L2 phones. In contrast, for adults, creating new categories of L2 phonemes is a more complex task, because their fully developed L1 system influences the assimilation of L2 sounds; however, according to Flege (1999), with the right L2 input this task is not impossible to achieve. For example, the formation of new phonological contrasts such as English /i:/ and /ɪ/ or /u:/ and /ʊ/, depends mainly on the L2 input received, opportunities for exposure to the target language, explicit phonetics instruction and the time dedicated to developing the set of new phonemic categories.

Different theoretical models have attempted to explain and describe the process that second language (experienced and naïve) learners undergo when learning non-native segments. These models suggest different results depending on the learners' previous linguistic experiences. The models referred to in this section are the Critical Period Hypothesis, the Speech Learning Model (and the revised version of it), the Perceptual Assimilation Model and the Second Language Linguistic Perception Model.

It is acknowledged that there are other models for L2 development (see e.g., Boersma 1998; Kuhl 1991); however, the most influential and potentially relevant ones for this thesis are set out below.

### **3.2.1 The Critical Period Hypothesis (CPH)**

The Critical Period Hypothesis (CPH) formulated by Lenneberg in 1967 proposed that there is a specific period when the brain must receive stimulation to continue with its normal function for acquiring a language with native-speaker proficiency, usually before puberty. Brain maturation leads to a left-hemispheric language specialisation (lateralisation); during this process, the brain begins to lose plasticity, narrowing the ability to acquire a language. That is, children must receive sufficient linguistic input to acquire their L1 normally and efficiently during the critical period, otherwise, full competence in their L1 is reduced or difficult to attain. However, while the CPH was initially formulated for L1 acquisition, it was extended to second language learning. The basic concepts remain the same, but when applied to L2, the expression ‘critical period’ has often been replaced by ‘sensitive period’, indicating a weakening of the position but nevertheless indicating that adult L2 learners are less sensitive to input than children. In other words, for an L2 learner achieving native-like pronunciation after puberty will be difficult to attain.

### **3.2.2 The Speech Learning Model (SLM)**

The Speech Learning Model (SLM) proposed by Flege (1995) introduces new hypotheses and essential principles in second language acquisition related to the production of segments by advanced adult speakers of an L2. The SLM states that most of the errors in the L2 are made due to an inaccurate perception of the target phonemes but predicts that adult learners whose L1 has a smaller vowel inventory than English (e.g., Spanish, Polish) will be capable, with time and practice, of recognising phonetic differences between L1 and L2 vowels. For the SLM, L1 and L2 sounds are connected through an ‘interlingual identification’ unconscious process, which is activated at the first exposure to the target language.

The model states that successful L2 learners (of all ages over a lifetime) will create new phonetic categories for the L2 vowels if the segments are perceived as new or unrelated to the speaker’s L1 vocalic inventory and, ultimately, will produce them in a new phonemic category representation. In other words, if the L2 learner can distinguish the differences between L1 and L2 vowels, new L2 vowel classifications can be created. On the other hand, if the L2 learner perceives L2 phonemes as comparable to the L1 segments an assimilation

process will occur whereby the vowels belonging to two phonetic sets in the L2 may be merged and produced in a similar way; new phonemic categories will not be created. Hence, the closer the cross-linguistic distance between L1 and L2 is perceived to be, the more remote the potential is of achieving the correct production of foreign vowels. According to the SLM, L2 speech learning is a slow process that requires intensive input from native speakers to succeed.

In support of the SLM, Lee et al. (2020) tested the impact of English familiarity on the perception of English vowels by L2 learners. In their study, Korean participants living in the United States (US) and the United Kingdom (UK) performed vowel identification tasks demonstrating that experience with the target language had a strongly positive effect on the way that L2 speakers perceived new tokens. This was found to be especially true for adults living in the UK. Overall, the findings suggested that the perception of some segments would be enhanced as learners gain familiarity with the L2 variety, corroborating the SLM position in which speech learning is an active process (even after puberty) which varies from individual to individual and that perception categories are applied later on in production.

### **3.2.3 Revised Speech Learning Model (SLM-r)**

The Revised Speech Learning Model (SLM-r) by Flege et al. (2021) retains some premises from the original version, but there are also some differences. The SLM-r maintains the idea that the reorganisation of phonetic systems occurs as a consequence of the L2 input received during a lifespan. However, in this revised model, the focus is shifted from advanced second language learners to early and late learners without any reference to their level of attainment. This is so for three essential reasons. First, the Critical Period Hypothesis presents a lack of consistency in explaining age-related learning differences. Second, the quality and quantity of L2 input play a key role in the production of target segments, which means that different speakers (of all ages) are exposed to diverse input. Third, the notion that advanced L2 learners will reach a 'final or plateau' level of attainment is no longer asserted, because it has not been empirically tested or verified for L2 speech learning.

This revised version of the model can be related to the results of a study conducted by Sancier & Fowler (1997) which shows a case study of L1 Portuguese - English adult speaker (late learner) who spent 2.5 months in Brazil and 4 months in the United States (switching

languages accordingly) in order to see if the time of exposure to one specific language would have an effect on the production of the voiceless stop consonants of the other language. The results showed that the production of the Portuguese voiceless stops, although always shorter than the English ones, were adapted while the participant was in the US, but, after several months in Brazil, the English VOT values drifted towards the Portuguese ones. According to the SLM-r, this can be explained by the fact that the voiceless stops in both languages were similar phones, such that the speaker combined the L1 and L2 categories allowing the change of the realisation of the L1 and L2 segments according to the language environment.

### **3.2.4 The Second Language Linguistic Perception Model (L2LP)**

The Second Language Linguistic Perception Model (L2LP) by Escudero (2005) is congruent with PAM and PAM-L2 but aims to explain the total process of L2 speech perception based on the individual learner. In this respect, the L2LP model states mainly how L1 is an influential factor in learning L2 sounds, then posits different learning scenarios which could impact on the progress of an L2 learner's perception. The sequential scenarios are presented as a) Initial Stage, b) Learning Task, c) Development Stage and d) End State.

a) *Initial Stage*: in this stage, L2 sounds are perceived as a single L1 class (categorised dimension) so that the L2 speaker needs to divide the L1 category or develop a new one (uncategorised dimension). For example, in the case of Spanish learners of English, vowel height (as indexed by the first formant frequency - F1) is an instance of a categorised dimension, whereas duration - which is not involved in phonemic contrasts in Spanish - represents an uncategorised one. Therefore, at this initial stage, Spanish speakers will perceive the English contrasting vowels /i:/ and /ɪ/ as the Spanish /i/, and /u:/ and /ʊ/ as the Spanish /u/ without implementing a systematic durational contrast because this perceptual parameter does not underpin the phonemic vowel contrast of Spanish.

b) *Learning Task*: this stage has perceptual and representational sub-tasks. They consist of making changes to, or creating new categories for, the already categorised dimensions, such as (in the case of Spanish speakers) the formation of a novel category to signal the duration of the English vowels and/or producing changes in their F1 values to classify, for example, English /i:/ and /ɪ/ vowels.

c) *Development Stage*: in this phase, distributional and gradual learning are in place. The first of these terms refers to the creation of a new uncategorised dimension such as vowel duration in the case of Spanish learners of English, while the second refers to the adaptation of the new L2 categories.

d) *End-State*: in this final state, the accurate perception of L2 sounds is not guaranteed. Thus, the quality of L2 input is a key factor to perceive the target sounds. The model agrees with the idea of previous ones as the perception of new sounds (L2) relies on the speakers' L1 system which will impact on the development of the new language. However, this model (in common with PAM, and PAM-L2) focuses mainly on the perception of sound contrasts by non-native speakers. Unlike the SLM, it is not centred on L2 segment production in respect of a specific non-native speaker group.

Overall, the above-mentioned models have hypothesised how L2 segments and vowel contrasts can be perceived and produced by non-native speakers. In addition, they attempt to further understand the processes and difficulties that L2 learners face when producing and perceiving L2 tokens.

The major components, points of overlap, similarities and differences between the L2 learning models set out above are somewhat convoluted and difficult to summarise in any concise way in continuous text. In view of this a summary table is presented in the next section.

### **3.2.5 Perceptual Assimilation Model (PAM and PAM-L2)**

In line with the SLM and SLM-r, the Perceptual Assimilation Model (PAM) (Best, 1995) and its variant the PAM-L2 (Best & Tyler, 2007), also postulate that discrimination of distinct L2 sounds depends on recognising how perceptually alike L1 and L2 segments are, such that the first language helps to construct the perception of new L2 phonemes. This model uses three main classification systems for non-native phones: categorised, uncategorised, and non-assimilated sounds. The classification *categorised* means that if new non-native sounds are similar to L1 phonemes, they can be considered either good or poor examples of the L1 sounds (the less similar the L1 and L2 segments are, the better the perceptual discrimination). *Uncategorised* refers to L2 segments that are not present in the speaker's L1 inventory; therefore, they are not assimilated into the L1, and finally, *non-assimilated category* refers to

L2 sounds which are extremely unusual for the L2 learner to the point of not being considered or identified by the L2 learner as a speech sound.

PAM takes into account phonetic and phonological points to explain how a native language can limit or benefit L2 phoneme perception, but, unlike Flege's SLM, it focuses mainly on perception and not production. In addition, PAM-L2 (an extended version of PAM) states that the quality and quantity of the L2 input is essential for language learning, and that the different experiences (or exposure) that non-native speakers will face in their life will shape their L2 learning outcome (e.g., bilingual vs L2 learners). One of the main differences between these models and the SLM is that the source of information which listeners use to discriminate between speech sounds is articulatory gestures for PAM and PAM-L2 but acoustic cues for SLM. These two models (PAM and PAM-L2) predict a learning process. However, they do not explicitly state how the progression works.

### **3.3 Summary**

The following Table presents a summary of the main principles of each model to illustrate their similarities and differences.

**Table 5**

*Main similarities and differences across models*

<b>Similarities/differences</b>	<b>CPH (Lenneberg, 1967)</b>	<b>SLM (Flege, 1995)</b>	<b>L2LP (Escudero, 2005)</b>	<b>PAM (Best &amp; Tyler, 2007)</b>
<b>Optimised learning during childhood</b>	✓	✓	✓	✓
<b>Emphasis on production</b>	X	✓	X	X
<b>Emphasis on perception</b>	X	X	✓	✓
<b>Emphasis on both</b>	✓	X	X	X
<b>Common L1/L2 phonological systems</b>	X	✓	X	✓
<b>Special importance placed on quality of L2 input</b>	X	✓	✓	✓
<b>Links level of proficiency possible to proximity of L1 and L2 phoneme systems</b>	X	✓	X	✓
<b>Proposes series of separate stages within learning process</b>	X	X	✓	X
<b>Gives insights into specific difficulties faced by Spanish learners of English</b>	X	✓	✓	✓

The above table is, of course, something of a simplification. It is fully acknowledged that in many cases the differences between the models are not categorical, but more ones of emphasis. However, the purpose of the table is not to set out the subtleties or nuances of the positions, but to cut through those and provide the reader with the principal differences and similarities at a glance.

In order to understand the present state of theorising about the learning of L2 phonetics and phonology, one needs to see the models in the context of their historical development from CPH to PAM. However, it will be clear from the descriptions above that some are more relevant to the present study than others. For example, those such as CPH and PAM, which mainly discuss the differential abilities of children and adults in attaining native-like phonological proficiency, cannot shed a great deal of light on the results of this study, as the subjects are all adults aged between 18 and 27 years. However, others such as SLM, which proposes proximity of L1 and L2 phoneme systems and disparities in how heavily populated the L1 and L2 phoneme inventories are, would appear to have greater relevance. This is because, as shown in Chapter 2, Spanish has a reduced vowel system compared with English, and the English vowel categories with which the present research is concerned are relatively near to, but slightly different from, those found in Spanish. Also, L2LP, which specifically addresses the issues around learning the English contrasts involving /i:/ and /ɪ/ and /u:/ and /ʊ/ for Spanish learners is particularly apposite. I return to the models in the discussion of the findings in Chapter 10.



## **Chapter 4. Data Collection**

### **4.1 Introduction**

The present research examines the progress of adult Spanish-speakers in developing productive contrasts between the English vowels /i:/ and /ɪ/, /ɪ/ and /e/ and /u:/ and /ʊ/ in respect of quality and duration. This chapter discusses the process of data collection for the project. It starts with a description of the participants, a sampling summary, recording procedures, and the stimuli used during the different data elicitation sessions. The sections dealing with these matters are followed by a description of the language background questionnaire (LBQ) which was implemented at the point of each recording session to gather information about everyday use of, and exposure to, English as well as personal extracurricular activities. The main purpose of the LBQ was to identify factors that might have an effect on developing proficiency in L2 speech production generally and specifically on the segmental contrasts in question.

### **4.2 Participants**

The participants were recruited based on the eligibility criteria for this study which involved the individual's language background (L1 Spanish), university enrolment status, English proficiency level, and nationality. They were contacted and invited to participate via email; following acceptance, a voluntary response sample was obtained.

#### **4.2.1 Language background and nationalities**

All of the participants were born, raised, and lived in a Spanish-speaking country before moving to the United Kingdom for their university enrolment. Their first language was Spanish, and they all had undergone formal education in Spanish in their home countries. For all the participants, English was learned as a foreign language in traditional classroom settings (2-4 hours a week), and none of them had ever lived for more than three weeks in an English-speaking country before moving to the UK.

The following 8 nations were represented in the sample: Mexico, Spain, Ecuador, Chile, Colombia, El Salvador, Perú, and Argentina. For the purpose of this research, the fact that speakers come from different countries and regions does not materially affect the production

of English segments, given that Spanish vowels only differ marginally, if at all, between varieties (Salcedo, 2010). Cross linguistic interference is therefore likely to remain uniform across the board.

#### **4.2.2 Enrolment status**

For this research, only participants who had enrolled as full-time students at the University of York (at undergraduate, masters, or doctoral level) were recruited. The purpose of this criterion was to be able to easily invite participants in for recording at several points during a year, and to ensure that everybody started their university experience at the same time – in October 2018. Thus, the subjects were all in their first year in the city of York and had arrived approximately at the same time.

#### **4.2.3 English level**

All of the subjects had taken an international English test as a prerequisite to admission to their academic course of study. The average result of the tests was 6.5, which according to the *Common European Framework of Reference for Languages: Learning, Teaching, Assessment* (2001) is equivalent to an independent user B2, which in general terms means that the L2 speaker can use the target language fluently in academic and social contexts without difficulty. None of the subjects involved had conditional admission, which means that no-one was required to take pre-sessional English summer courses at the University in order to enter their programme.

### **4.3 Sample**

The initial sample consisted of 43 speakers: 21 females and 22 males. Of those, three individuals were excluded from the analysis (1 female, 2 males) for failing to complete the full series of recordings.

The average age of the remaining (20 female / 20 male) participants was 27 years ( $M = 27.3$ ,  $SD = 2.81$ ). The following countries were represented in the final sample: Mexico (12), Spain (13), Ecuador (5), Chile (4), Colombia (3), El Salvador (1), Perú (1), and Argentina (1).

#### **4.4 Recording procedures**

The individual recordings were made in a sound treated recording studio. The equipment used for the recordings was the following:

- (i) Microphone - dpA 4066 head-worn omnidirectional microphone; frequency response 20 Hz to 20 kHz, 3 dB soft boost at 8-20 kHz;
- (ii) Microphone amplifier and Analog to digital Converter - Behringer Xenyx U1204 mixing desk (USB output).
- (iii) Audio Software - Adobe Audition CC 2019;
- (iv) Recording set up running on Apple iMac MacOS 10.13.6 High Sierra.

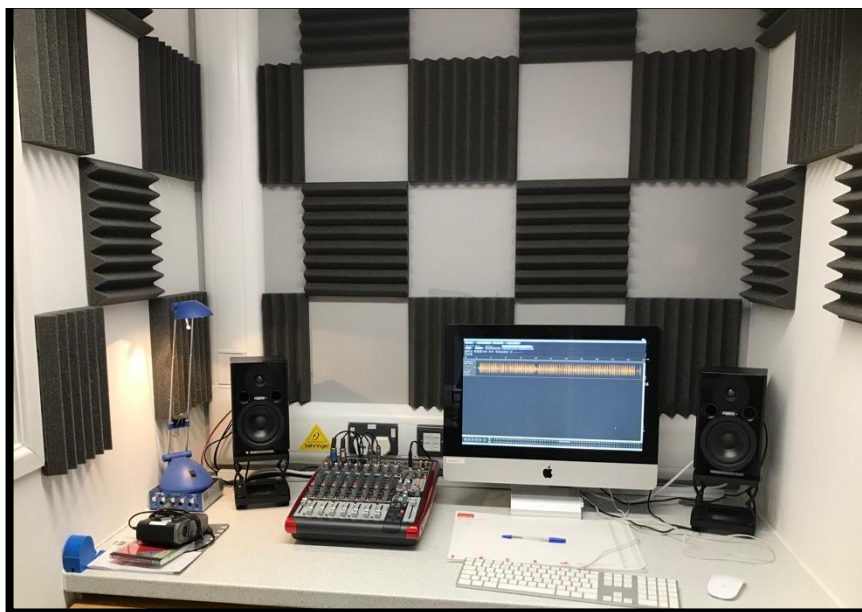
The recordings were made with a sampling rate of 44.1 kHz and bit depth of 16 and saved in .wav format. All data were collected without any EQ or filtering applied. Twelve dB of headroom was allowed during the recording process to avoid overloading or clipping of the signal.

Participants were asked to introduce themselves, describe pictures, read passages and a list of 60 words, self-paced and in a single session without a break. The same procedure was applied for each session at three time points in the year, using different stimuli.

The following Figure shows the recording equipment.

## Figure 6

*Recording equipment used during data collection*



### 4.5 Data management

Prior to data collection, and in line with the University of York and the Language and Linguistic Science Department ethics guidelines, participants received an information sheet (see Appendix A). They gave informed consent and were made aware that they could withdraw at any point during data collection (see Appendix B).

Before each recording session, participants were given a code number, which replaced their name, to ensure all data were anonymised. After the recordings, the data were stored on a secure University of York server and backed up on a password-protected, encrypted 1 TB external hard drive. Language background questionnaires and consent forms were safely stored in a locked drawer. Data analysis and discussion in this thesis refers to participants only by their code number, thus guaranteeing participants full anonymity.

## 4.6 Method

The following section explains the methods used to elicit and record participants' speech followed by a description of the language background questionnaire. Data analysis methods are discussed in the following chapter.

### 4.6.1 Participant recordings

#### a) Stimuli

Three different production tasks were recorded at three different times during the year. The first session was recorded one month after arriving at University (November 2018); the second session was recorded five months after the first one (April 2019), and the final session was recorded at the end of the academic year (September 2019). The recording sessions were self-paced and conducted without a break or practice beforehand.

The *production task* contained four elements (see Appendix C). Participants were recorded in each session producing:

- (i) spontaneous unscripted self-introduction conversation;
- (ii) reading of two phonetically balanced passages from a pool of six passages: *The North Wind and the Sun* (International Phonetic Association, 1999), *My Grandfather Passage* (Van Riper, 1963), *The Story of Arthur the Rat* (Abercrombie, 1964); *The Boy who cried Wolf* (adapted from Aesop's Fables), *The Caterpillar* (Patel et al., 2013), and *The Rainbow Passage* (Fairbanks, 1960);
- (iii) engaging in a picture description task (three pictures each time);
- (iv) reading from a word list (60 words).

All the elements involved in the recording task were pre-tested to ensure their viability. The pilot study involved six Spanish speakers, all graduate students between 27 and 29 years old, from Argentina, Chile, and Mexico. None of them was a participant in the main study. The pre-test (dis)confirmed that the practical and formal aspects of each item were clear and doable within a reasonable amount of time. After the pilot study, the materials were revised and changes were made; specifically, different pictures were chosen for Time 2, and two

words from the word list for Time 1 were changed due to uncertainties in their pronunciation (see Appendix C for full list of reading materials).

### 1. Spontaneous unscripted self-introduction conversation:

In this part of the recording session, the participants introduced themselves by giving information about their nationality, studies, and the activities they had carried out the previous day, among other informal topics. The purpose of this unscripted talk was to create a relaxed and comfortable environment, allowing the speaker to get used to the microphone and the fact of being recorded. Also, by creating this English self-introduction, the subjects had the opportunity to briefly speak English before starting the reading task, and it thus worked as a 'warm-up'.

### 2. Reading two passages:

Phonetically balanced passages are commonly used to elicit speech data. The meaning of 'phonetically balanced' refers to the passage having all the phonemes used in the language in question represented in the passage, the number of tokens of each reflecting their frequency of occurrence in contemporary English (Gibbon et al., 1997). Two different passages were selected in every session. Since participants were proficient English users, they read the passages within 5 minutes.

### 3. Picture description:

This task involved asking participants to describe a set of three different colour pictures. The aim was to enable them to speak in a partly controlled but non-scripted manner. The images were taken from different versions of the Test of English as a Foreign Language (TOEFL). Prior to using each set of photos, the images were pre-tested to ensure that they were clear and had elements that could readily be described.

#### 4. Word list:

Word lists, eliciting citation forms, are well used in experimental studies to elicit speech data (see Polka, 1995; Strange et al., 2007). Some of the advantages of using this kind of material is that the phonological context of a vowel can be controlled to avoid the effects of connected speech processes; the influence of factors such as context and reading fluency are reduced, therefore, takes the speaker nearer to competence than other speech elicitation methods. Isolated words tend to be produced with more articulatory distinction, slower pace, and pauses than e.g., spontaneous speech. In other words, data from citation forms can be the most canonical full form of a word. (Ladefoged & Johnson, 2010; Strange et al., 2007).

For this study, the word list comprised 60 monosyllabic English words. The purpose of using only monosyllabic words was to avoid problems such as stress misplacement that could affect the production of the vowels targeted in this research. The online program *English Lexicon Project* (Balota et al., 2007) was used to generate high frequency words to avoid doubts in pronunciation while reading, which may occur with less frequent and therefore potentially unfamiliar words. Each word in the list had a CVC structure and comprised between three to five orthographic letters in order to limit the set of contexts and obtain controlled data. Word order was randomised before presenting them to participants using the programme Random.org (Haarh, 1998).

All the recorded tasks yielded speech data, and the analysis was undertaken on the data obtained from the different word lists and read passages. The picture descriptions and free conversations data were not analysed in this research due to the lack of consistency in the length of the responses; however, the material will be used in a follow up study.

Figure 7 illustrates the recording process.

## Figures 7

*Sample of a recording session*



### 4.6.2 Written Language Background Questionnaire

Questionnaires are often used in linguistic research to gather evidence on, for example, motivation, attitudes and previous experience with a language. A fixed questionnaire - as opposed to, say, a free interview approach, provides standardisation and control over information elicitation across subjects (Selinger & Shohamy, 1989).

Prior to administering the questionnaire, it was piloted on five people, all graduate students between 27 and 34 years old from the UK, Mexico and Spain, to ensure clarity of the questions, check formal aspects (spelling, grammar etc.), and assess the approximate time required to complete the full task. After pilot testing the printed version of the form, it was revised, and changes were made to improve the quality of the data being collected. Hence, three additional questions were added to the final questionnaire only, which was applied at Time 3. These items related to a self-assessment or self-reflection on the English and Spanish language spoken during the year( see Appendix D, questions 26, 27, 28).

The questionnaire mainly comprised structured questions, and answers were required to be expressed on a 1-10-point Likert scale. Other questions were included to obtain background



information on matters such as age, nationality, languages spoken, etc. There was also a final, open-ended question that allowed participants to add any extra information they considered important. Another main purpose of asking structured questions was to gather information on language usage and track possible changes that could have occurred during the year with regard to their L1 or L2 exposure and use (see Appendix D for the full questionnaire).

In order to avoid the low-response rate sometimes observed in respect of online questionnaires, all participants received a printed version of the questionnaire after having completed the recording tasks, which they filled in 'on the spot'. This was done every time subjects were recorded (three times in one year).

The questionnaire was divided into four main parts:

- I) Background information
- II) Attitudes towards English
- III) Opportunities to develop the English language
- IV) Circumstances of exposure to English

The background information section was used to gather personal information such as age, country of origin, field of study, English test score etc. The attitudes section dealt with personal attitudes towards the use of English and Spanish, questions here were related to accent varieties, importance or otherwise of maintaining L1 accent, and self-evaluation of English skills, among other matters. The English opportunities section included questions associated with personal chances to interact with native speakers in academic and non-academic contexts, such as supervision or tutorial time, extracurricular activities, and on and off campus work. The last section of the questionnaire concerned the different circumstances that allow participants to interact with native English speakers, and had questions focusing on living arrangements (house-sharing), friendship developments, English and Spanish language use, and so forth.

Together, the four sections of the questionnaire provided potentially important information about participants' exposure to and use of English. This material has proved to be valuable in interpreting the patterns found in the speech production data. (see Chapter 9)

## 4.7 Pre-analysis procedures

Prior to the final analyses of the data gathered from the different recorded materials, two pre-analyses were conducted to prepare the final data: 1) normalisation of the raw data; 2) independent t-test for gender variable.

### 4.7.1 Normalisation

Normalisation measures have been developed and used to reduce differences in individuals' formant values, caused by anatomical and physiological variation, such as vocal tract length which may be associated with speaker gender (Strycharczuk & Scobbie, 2017; Thomas & Kendall, 2007). For instance, female speakers commonly produce higher frequency formants than male speakers because women's vocal tracts are shorter than men's making the comparison of vowel quality measurements across speakers problematic. According to researchers, vowel normalisation techniques have four main goals which can vary according to the specific research necessities. These are to: a) eliminate physiological differences among speakers b) preserve sociolinguistic features in vowel quality, c) retain phonological vowel distinctions, and d) model the human perceptual system. (Thomas & Kendall, 2007; Watt et al., 2010). Different normalisation methods can be implemented in the vowel normalisation software (Norm) by Thomas & Kendall (2007) such as Bark Difference Metric, Labov ANAE, Lobanov, Nearey and Watt & Fabricius methods.

The vowel extrinsic Lobanov method was used in this research as a pre-analytic procedure, because it compares formant values of different vowels spoken by a speaker with the main advantage of reducing physiological variations in formant values while keeping the sociolinguistic ones (Clopper, 2009; Adank et al., 2004). The formula applied in this method in Norm is the following:

$$F_n[v]_N = \frac{F_n[v] - MEAN}{S_n}$$

$F_n[v]$  represents the normalised value of a vowel (i.e., F1/F2).  $MEAN_n$  is the average for the formant (F1/F2) produced by the speaker and  $S_n$  is the standard deviation of the specific formant analysed (Thomas & Kendall, 2007).

### **4.7.2 Gender variable**

Preliminary analyses with normalised data were conducted at an early stage in the research to see if gender was a significant variable to include in this thesis. Independent t-tests were conducted for formant-gender, and duration-gender effects over the three time points for all the vowels examined. The results (after normalisation) showed no statistically significant effect for gender; therefore, this variable was excluded.

### **4.8 Summary**

The purpose of this chapter was to describe the instruments and procedures employed for data collection, starting with a description of the participant recruitment and recording techniques, and ending with the materials used for speech recordings and the format of the language questionnaire. Also, the description of the pre-data analyses conducted on all words used throughout this research was explained.

The speech data obtained from the word lists, the phonetically balanced passages and the information from the written language background questionnaire, in addition to further data obtained from a perception test by native English speakers (which is discussed later in Chapter 7) form the material analysed in this thesis.

## **Chapter 5. Marking of English Vowel Contrasts over a Year: Overall group performance**

### **5.1 Introduction**

The purpose of this chapter is to provide a general view of the production of the English vowel pairs contrast /i:/ and /ɪ/, /ɪ/ and /e/ and /u:/ and /ʊ/ with respect to both spectral and temporal features by adult Spanish speakers (as a group) in a citation style of speech at three points in time. It lays the groundwork for the later Chapters 6 and 8, which analyse in detail variations in progress across time at the individual level, and the possible factors influencing the differences among speakers.

#### **5.1.1 Structure of the chapter**

The sections of the chapter are as follows:

- background information
- acoustic measurements (formant and duration analyses) taken over a year at three points in time
- results for both spectral and temporal features
- summary

### **5.2 Background information**

It is commonly known that speakers can modify their speech according to the context (see for example, Coupland et al., 1988; Drljača, 2017; Muir et al., 2017). When speakers wish to be more intelligible, both native and non-native English speakers are reported to tend towards slower, louder, and a more highly articulated style known as ‘clear speech’ (Smiljanić & Bradlow, 2005). Therefore, the challenging task in analysing speech changes is to control the context in which speakers are tested and regulate their speech styles (Labov, 1972). Word lists are a frequently-used tool for eliciting speech data. One advantage of using them is that phonological context can be controlled to avoid connected speech processes such as vowel reduction, coarticulation effects, and prosodic features that could affect the final pronunciation outcome (Fogerty & Humes, 2012; Shattuck-Hufnagel & Turk, 1996). Another advantage is that, in Chomskian terms (Chomsky, 1965), they may allow the researcher to get nearer to the underlying phonological/phonetic competence of a speaker than do other

elicitation methods. In wholly unscripted talk, the speaker is likely to be focussing on, among other things, the content of his/her message, and in reading from a scripted passage, the speaker may well be attempting to focus on the prosodic fluency of his/her delivery. With a word list, the possible influence of these factors such as stress, intonation and rhythm is minimised. Notwithstanding the premise that one can never access competence – a speaker's unconscious underlying knowledge of a system – except by observing performance, one might argue that the word list, where the emphasis is on clear pronunciation of isolated words, reduces performance factors as much as can be achieved. It therefore takes one nearer to competence than do other speech elicitation techniques.

Word list elicitation has a long pedigree in respect of vowel studies, including the vowels under examination here. While the results of some studies (e.g., Smiljanić and Bradlow, 2008) found that temporal differences for tense and lax vowels in connected speech were not dissimilar to clear speech, most give strong support to the position outlined above. A study by Leung et al. (2016), for example, assessed temporal and spectral differences between the production of members of the English vowel pairs /i:/-/ɪ/, /a:/-/ʌ/, and /u:/-/ʊ/ in clear speech. Six isolated words (in KVD context e.g., 'keyed', 'kid', 'cad') were read by 18 native speakers of Canadian English. Differences in the temporal and spectral dimensions were found. When speakers produced words in a more articulated manner, the tense vowels /i:/, /a:/, and /u:/ increased in duration, while the lax vowels /ɪ/, /ʌ/, and /ʊ/ showed more formant changes. That is, the vowel pairs examined (in clear-citation form style) went through different dimensional changes. Another study, conducted by Smiljanić and Bradlow (2005), which tested the effect of clear speech in English and Croatian, found that word list productions increased the F1 and F2 vowels' values, raising the degree of separation between contrasting vowels (in both languages).

### **5.2.1 The acoustics of vowels**

### **5.2.2 Formants**

Fant's (1970) source filter model describes speech sounds and their production; in the model the quality of a vowel sound is determined by the peaks of the amplitude of the filtered harmonics. These spectral peaks in the frequency spectrum - the formants - are numbered starting from the lowest to the highest resonant frequency, the lowest being F1, the subsequent peak F2, then F3, and so on. Researchers consistently use the two lowest formants (F1 and F2) to describe and define the quality of a vowel sound in an unbiased way (Pols et al., 1969; Watt & Fabricius, 2002; Wright & Nichols, 2015). As noted in Chapter 4, F1 and F2 values are used to locate pronunciations within vowel space. Thus, F1 and F2 values are measured and referred to throughout this research.

### **5.2.3 Vowel duration**

Duration is another characteristic (besides quality) that speakers may use to distinguish vowels phonemically from each other. When duration is a contrastive feature (primary or secondary) as in the English language, the relative duration of a vowel is described and compared to similar vowels within the same vowel inventory (Hillenbrand et al., 2000; Wells, 1962). The length of a vowel can be affected by different factors such as the consonant context, rate of speech, intonation and stress (Van Leussen et al., 2011). As explained in Chapters 2 and 4, Spanish and English differ in the significance of vowel duration, and further analysis is provided in the following sections.

With regard to English vowels, the terms long/short and tense/lax vowels are used throughout this research interchangeably, in accordance with which terms authors being cited choose to use.

### 5.3 The English Vowel Contrasts: /i:/ and /ɪ/; /ɪ/ and /e/ and /u:/ and /ʊ/

It is now almost universally recognised that the members of these three vowel pairs may be phonemically differentiated by both quality and duration. This recognition on behalf of phoneticians and phonologists was evident in a debate concerning the symbols to be used to represent them that took place among members of the International Phonetic Association and others, concentrated mainly in the early 1960s. The preference of Daniel Jones, evident in his symbolisation in the various volumes of the English Pronouncing Dictionary (EPD) that he produced (see e.g., Jones, 1917, 1926), was not to mark quality. Rather, in order to avoid proliferation of symbols, he used the two single symbols with and without length marks /i:/ and /ɪ/ and /u:/ and /ʊ/ to make the distinctions. This approach has been termed the *quantitative* position (see Wells, 2001). A number of phoneticians were dissatisfied with this arrangement and considered quality rather than quantity to be important and proposed that a different basic symbol should be used for each member of the two pairs /i/ and /ɪ/, /u/ and /ʊ/ the *qualitative* position (see e.g., Abercrombie, 1964). The argument was finally resolved when Gimson superseded Jones in respect of EPD production. Gimson (1962) adopted the almost universally accepted notations used in this thesis, which mark *both* quantity *and* quality /i:/ and /ɪ/, /u:/ and /ʊ/ (the *quantitative-qualitative* position).

However, irrespective of notational preferences, protagonists of both camps accepted the relevance of both dimensions in the phonemic contrasts, and this is supported by empirical studies of vowel perception such as those undertaken by Escudero & Boersma (2004) and Flege et al., (1997). Chapter 6 contains a discussion of the perceptual significance of quality versus quantity.

It should be noted that continual changes in the phonetic realisations of English segments make the acoustic data that has been published for RP/SSBE vowels difficult to fully apply to current pronunciation trends. However, in recent years studies have been carried out to complement the already existing data from Wells (1982), Henton (1983) and Bauer (1985) among others, and, although they may still be considered somewhat dated, they nevertheless provide the best available empirical data for vowel quality (see e.g., Bjelaković, 2016; Deterding, 1997; Fabricius, 2007; Ferragne & Pellegrino, 2010) and quantity values (see e.g., Ali, 2013; Wells, 1962; Wiik, 1965).

In this research, the target values for quality and duration towards which participants might be expected to be shifting if they are adapting to British English norms were taken from two

main sources. The quality reference values are from the data published by Deterding (1990, as cited in Deterding, 1997) which included 8 male RP speakers reading citation words. The relative duration values are from Wells (1962) who included 25 male RP speakers; while it is recognised this is an older study, nevertheless, is still being used as reference data in more recent studies (e.g., Ali, 2013) and no more up-to-date studies (with more than 25 RP speakers) were found at this point. In any case, as it will be seen later in the results of this chapter and the following Chapters 6 and 8, the RP duration values turned out to be irrelevant since the Spanish participants tended not to make any length distinction.

The following are the reference values:

**Table 6**

*Quality reference values*

<b>V</b>	<b>F1</b>	<b>F2</b>
/i:/	280	2221
/ɪ/	380	1960
/u:/	302	1130
/ʊ/	414	1051
/e/	560	1797

Deterding (1990 as cited in Deterding, 1997) citation words

**Table 7**

*Duration reference values*

**Duration**

Short vowels: 160 ms (.16sec) SD: .03 sec

Long vowels: 300 ms (.30sec) SD: .08sec

Wells (1962)

As stated in the Introduction chapter, this research aims to examine the progress of adult Spanish-speakers in distinguishing between English vowels /i:/ and /ɪ/, /ɪ/ and /e/ and /u:/ and /ʊ/. A strong rationale for a focus on these vowels is that it minimises the problem of which



variety of English the participants might be adapting to, because – even though they are going to be exposed to a range of different varieties at the University and outside of it – the short vowels /ɪ/ and /ʊ/ and the ones absent in Spanish are subject to very little regional and social variation in British English (French et al., 2008; Wells, 1982).

One might expect, as a consequence of immersion in spoken English, to find a development of the /i:/-/ɪ/ contrast to be marked by the movement of /ɪ/ and /ʊ/ to more open positions over the course of the year that the learners spent in England. However, if this were to occur with /ɪ/, it would begin to converge with native Spanish pronunciations of /e/, which are much closer than contemporary native English realisations (see Hickey, 2018). In order to maintain or create an /ɪ/-/e/ contrast, therefore, learners' pronunciations of /e/ would need also to change, i.e., to become more open and nearer to contemporary native English norms, i.e., around cardinal vowel number 3 [ɛ]. Such a change might, in any case, be expected, simply because of exposure to native English /e/, but the point is that there are two potential factors here, each pulling in the same direction, one systemic (a 'push' effect - see Lubowicz, 2011), and the other a natural, accommodatory gravitation towards the present-day /e/ target, as a result of exposure. If, owing to these factors, /e/ were to move towards a native English type quality, one would expect this to be mainly marked by a raising of F1 across the three recording time points to somewhere around 500 - 600 Hz (see Table 6 above).

It is acknowledged that the relative functional loads of the vowel pairs examined are unequal. The /i:/-/ɪ/ and /ɪ/-/e/ contrasts carry a relatively high functional load, i.e., there are many minimal pairs of English words that are distinguished by the occurrence of one versus the other member of the pairs. Failure to distinguish productively between members of these pairs could therefore lead to frequent occasions of misunderstanding on the part of listeners. In view of this, there may be a strong incentive for non-native speakers of English to master the distinctions. By contrast, /u:/-/ʊ/ carry a low functional load, i.e., this contrast is used to distinguish between few words. Failure to make a clear distinction between members of this pair would therefore be unlikely to result in frequent misunderstandings.

The inclusion of both low and high functional load pairs in the study allows one to gain some provisional insight into the relative weighting of factors responsible for participants changing towards native English norms. A clearly larger change in respect of the front pairs over the back pair would argue for the ascendancy of communicative needs – specifically, the need to

be understood. A roughly equivalent rate of change for the back and front pairs might suggest that the wish to sound proficient in English was at least equal to the communicative need impetus. In other words, one can make high rates of progress even if a strong communicative need factor is not present, if one is sufficiently motivated to achieve native-like pronunciation *per se*.

Another factor relevant to disparate functional load and these issues is that the low load of the back pair will mean that the learners will have only infrequently been exposed to meaningful exemplars of the distinction over the course of their year in York. If learners have achieved a high rate of back pair separation despite this, one could postulate that this indicates a high level of attention to the details of the speech of native English speakers coupled with a high degree of motivation to move towards their patterns.

The following sections present the analyses and results of the vowels produced over a year at three points in time by adult Spanish speakers (as a group) in a citation style of speech, i.e., from the word list data.

#### **5.4 Analysis**

The analyses presented below were conducted with reference to both spectral (vowel quality) and temporal (duration) features. Concerning duration, the analysis was made between short and long vowels in order to observe production differences, therefore, the front pair /i:/ and /ɪ/ and the back pair /u:/ and /ʊ/ were measured and compared (the /e/ is excluded from the duration analysis throughout this thesis, because it is a short vowel as well as /ɪ/).

Formant analysis is presented first, followed by the duration measurements and the statistical analysis.

As explained in Chapter 4, in order to control for physiological differences across speakers, all formant values were normalised, and female and male data were combined.

### 5.4.1 Formant analysis

For the purpose of the study 15 monosyllabic words containing the target vowels /i:/ and /i/, /ɪ/ and /e/ and /u:/and /ʊ/ were selected from the word list read across the three time points. A total of 1,800 tokens (15 words x 3 times x 40 participants) were obtained and analysed.

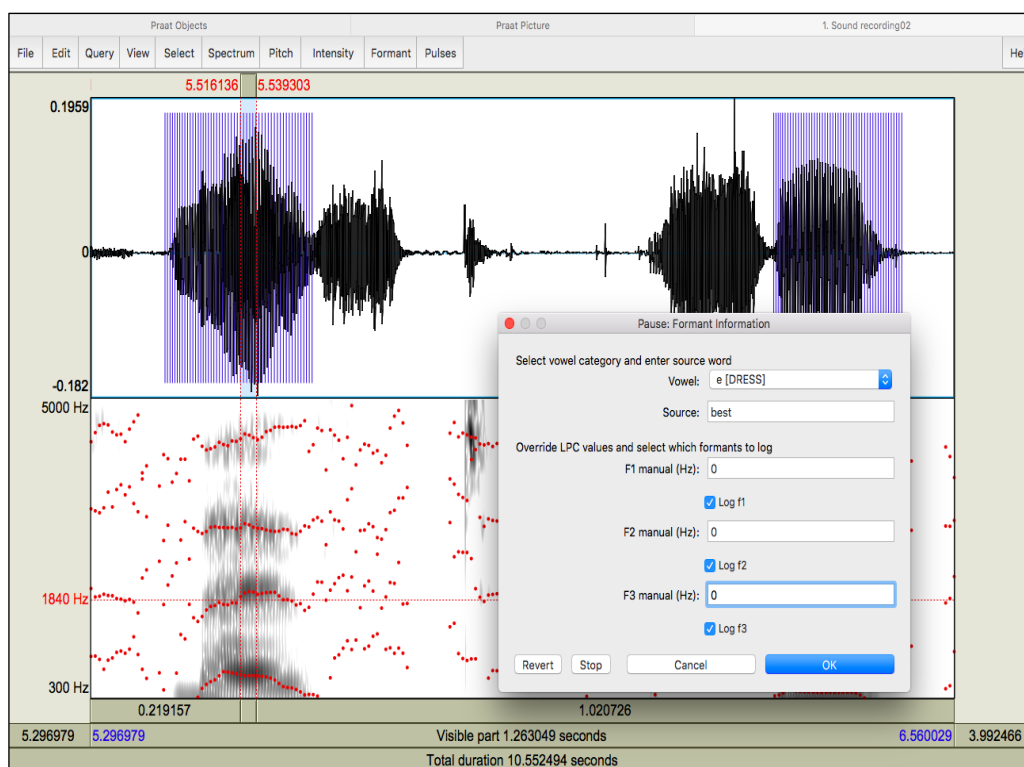
Words with palatal onset /j/ were not included in order to avoid fronting as a co-articulatory effect, as opposed to an adjustment to newly emerged L2 fronting norms (see Chapter 6), for /u:/ and /ʊ/. Similarly, no words with initial /w/ were included so as to avoid co-articulatory F2 reduction effects arising from the velarity and the lip/rounding of the initial consonant on both contrast pairs.

To obtain the F1 and F2 values, the words with the target vowels were first isolated from the read list recordings using version 3.0.2 of Audacity ® recording and editing software (2021), and then saved as .wav files. Formant measurements were manually extracted at approximately the midpoint of each vowel using Praat software version 6.0.49 with settings of max formants 5500 (Hz) and 4.5 as number of formants (with adjustment when needed) (Boersma & Weenink, 2019). Finally, the raw values of F1 and F2 were normalised as described in the previous chapter to obtain the speaker means per vowel.

The following figure is an example of a word extracted from the word list where F1 and F2 were measured with Praat.

**Figure 8**

*Sample of the analysed word.* The word uttered in this extract is the word ‘best’. The menu on the right was used to log the formants, vowel and source for retrieval and analysis.



## 5.4.2 Duration analysis

The occurrence of length distinctions between /i:/ and /ɪ/, and /u:/ and /ʊ/ can only be considered in one phonological context, namely in syllables that close with a voiced consonant. As explained in Chapter 2, the ‘short’ members of each pair do not occur in open syllables, and in syllables closed by a voiceless consonant the contrasts tend to be neutralised on the duration dimension because the ‘long’ members of both pairs are produced short, i.e., in a way similar to the short ones.

However, at Time 1<sup>3</sup> the duration of /i:/ was also examined where the syllable was closed by a voiceless consonant, because if durational distinction occurred under this circumstance, it

<sup>3</sup> Only T1 was examined for /i:/ closed by voiceless consonants because it was the starting point for the speakers, and it was taken as a reference and extra information.

might suggest that participants over-extended the syllable closed by voiced consonant length rule.

The duration of the two sets of vowels contrasts /i:/and /ɪ/, /u:/and /ʊ/ (1,440 tokens in total, 12v x 3Ts x 40sp) was measured in seconds using a Praat script, hand corrected and finally, converted into milliseconds (ms). Initial segmentation (in the form of TextGrids) was manually created to facilitate the extraction of the duration values. The vowel measurements were taken from the onset of voicing in the vowel to the offset of each vowel in establishing the speaker means per vowel.

### **5.4.3 Statistical analyses**

a) Formant analysis: A statistical Repeated Measures Analysis of Variance (ANOVA) was conducted separately for each vowel to detect differences in vowel quality across Time 1 (T1), Time 2 (T2), and Time 3 (T3). The within-subject effects for the ANOVA were time (3), formant (2), time \* formant. The interactions between time and formants were followed by post-hoc analyses: paired-sample t-tests ( $p < .05$ ) for which all significant results are Bonferroni corrected.

b) Duration analysis: In order to determine statistical significance in the quantity of vowels across Time 1 (T1), Time 2 (T2), and Time 3 (T3), the mean duration of the four vowels /u:/, /ʊ/, /i:/, /ɪ/ and /e/at each recording session was used in a Repeated Measures ANOVA. The within-subject effects for the ANOVA were time (3), vowel (4) and time \* vowel. Interactions between time and vowel, and distance between vowel contrast and time were followed by post-hoc analyses: paired-sample t-tests ( $p < .05$ ) for which all significant results are Bonferroni corrected.

## 5.5 Results

### 5.5.1 Formants

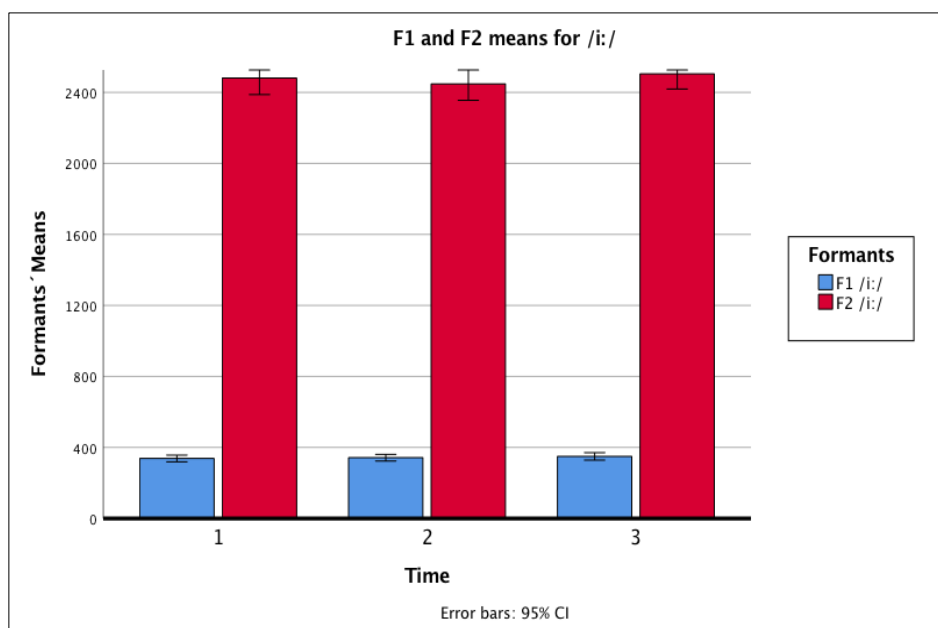
To obtain a general view of the performance of the group as a whole, the following results obtained for F1 and F2 were measured in the production of the English vowels /i:/ and /ɪ/, /ɪ/ and /e/, /u:/and /ʊ/, and compared at T1, T2, and T3 in order to see the effect of time over the production of both formants (the results are presented independently per vowel).

#### a) Vowel /i:/

The mean scores of F1 and F2 at three different time points were analysed in a Repeated Measures ANOVA. The results show significant differences in the vowel production over the time points [ $F(2,78) = 6.505, p < .002$ ], and non-significant effects in the interaction between time - formants [ $F(2,78) = 2.670, p > .076$ ]. Post-hoc analysis revealed a trend in times - F1 interactions but did not reach the level of significance. The outcomes for the interaction for times and F2 showed the same non-significant trend. The following Figure provides a visual outcome for the overall results.

**Figure 9**

*F1-F2 means for /i:/*



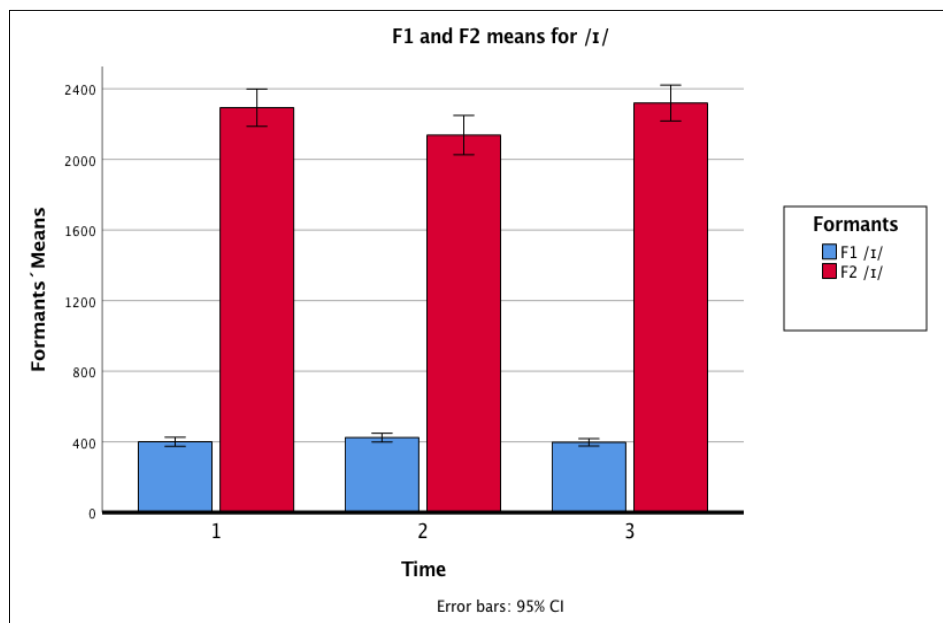
Note: Mean across times F1:343, F2:2477

## b) Vowel /ɪ/

The mean F1 and F2 of the vowel /ɪ/ at three time points was calculated, and a Repeated Measures ANOVA was performed. Greenhouse-Geisser correction was applied, showing significant differences in vowel production over time points [F (1.489, 58.087) = 17.173,  $p < .001$ ], and between time and formant interaction [F (1.593, 62.109) = 21.255,  $p < .001$ ]. Post-hoc tests indicated significant interactions between T1-T2-F1  $p < .018$ , and T2-T3-F1  $p < .036$ ; however, the interactions between T1-T3-F1 were not significant. Results for time and F2 showed significant differences between T1-T2-F2, and T2-T3-F2 at  $p < .001$ . However, the interactions between T1-T3-F2 were not significant (Figure 10).

**Figure 10**

*F1-F2 means for /ɪ/*



Note: Mean across times F1:407, F2:2250

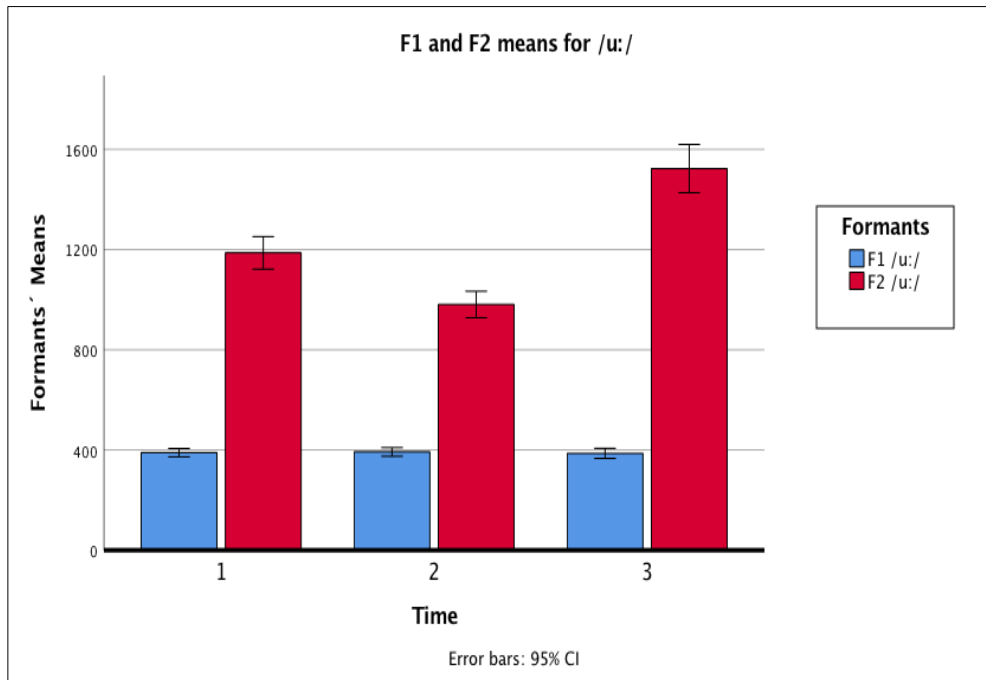
## c) Vowel /u:/

A Repeated Measures ANOVA with Greenhouse-Geisser correction showed that the mean of /u:/ significantly differed between the time points [F (1.291, 50.335) = 72.026,  $p < 0.001$ ]. Regarding time and formant interaction, results showed a significant effect as well. [F (1.250, 48.762) = 84.955,  $p < 0.001$ ]. Post-hoc analyses between time points and formants revealed non-significant differences between times and F1. However, statistically significant

interactions between times and F2 were found at  $p < 0.001$ . The next Figure illustrates the results.

### Figure 11

*F1-F2 means for /u:/*



Note: Mean across times F1:389, F2:1230

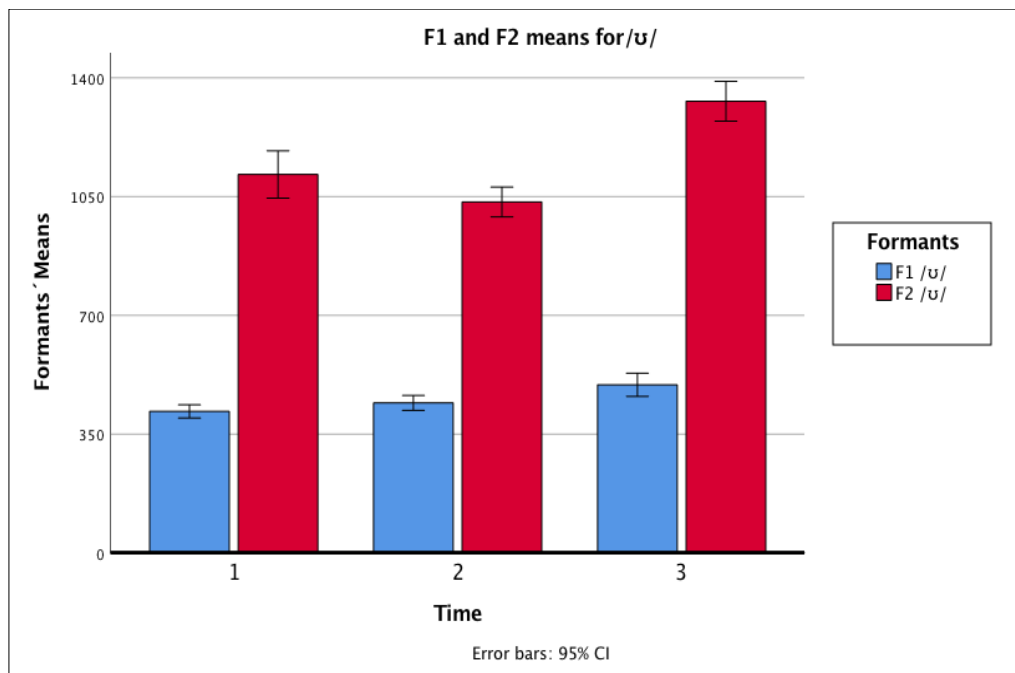
#### d) Vowel /o/

A Repeated Measures ANOVA was conducted, obtaining significant differences in the vowel production over the three time points [ $F(2, 78) = 51.632, p < .001$ ], and significant effects in the interaction time-formant [ $F(2, 78) = 50.172, p < .001$ ]. Post-hoc tests indicated significant results in the interaction of T1-T3-F1, and T2-T3-F1 at  $p < .001$ ; however, the difference between T1-T2-F1 did not reach statistical significance. Regarding F2, significant interactions were found in T1-T3-F2, and T2-T3-F2 at  $p < .001$ ; however, the interaction between T1-T2-F2 revealed non-significant effects. (Figure 12).



**Figure 12**

*F1-F2 means for /ʊ/*



Note: Mean across times F1:451, F2:1160

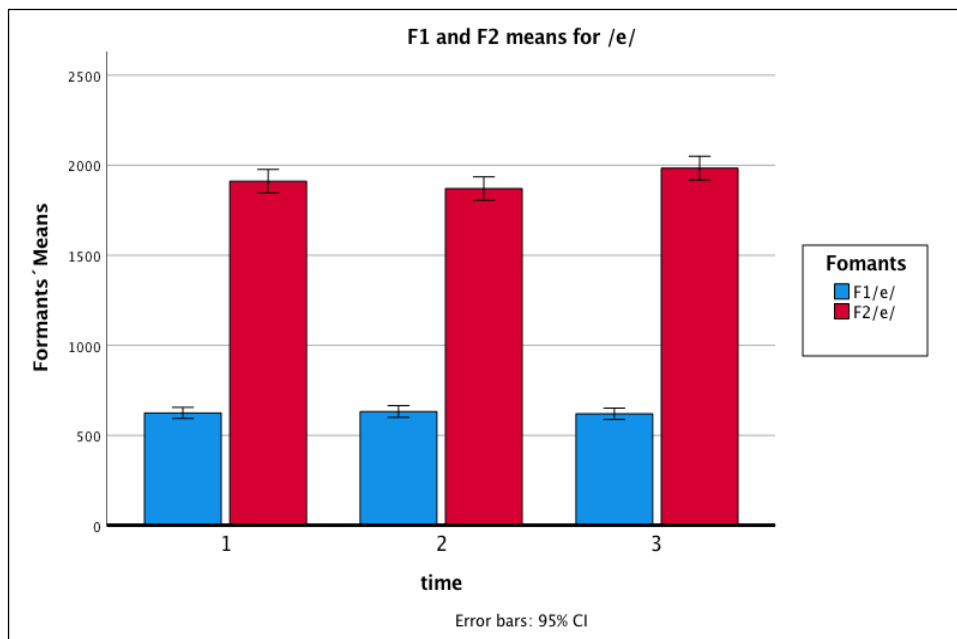
**e) Vowel /e/**

The mean scores of F1 and F2 at three different time points were analysed in a Repeated Measures ANOVA. The results show significant differences in the vowel production over three time points [ $F(2,78) = 17.3226, p < .001$ ], and significant effects in the interaction between time - formants [ $F(2,78) = 12.580, p < .001$ ]. Post-hoc tests indicated that results in the interaction of T1-T2-F1, and T1-T3-F1 were not significant. However, the interaction of T2-T3-F1 was significant at  $p < .001$ . Regarding F2, significant interactions were found in T1-T2-F2 at  $p < .015$  and the interactions T2-T3-F2 and T1-T3-F2 at  $p < .001$ .

The following Figure provides a visual outcome for the overall results.

**Figure 13**

*F1-F2 means for /e/*



Note: Mean across times F1:625, F2:1922

The following Table displays the results of the ANOVA and Post-hoc- test as a summary.

**Table 8***Summary of the F1 and F2 results*

Vowel	Time (x 3)	Time * Formants	Post-hoc Test F1	Posthoc Test F2
/i:/	$p < .002$	non- sig	non-sig	non-sig
/ɪ/	$p < .001$	$p < .001$	T1-T2 F1 $p < .018$	T1-T2 F2 $p < .001$
			T2-T3 F1 $p < .036$	T2-T3 F2 $p < .001$
			T1-T3 F1 non-sig	T1-T3 F2 non-sig
/ʊ/	$p < .001$	$p < .001$	T1-T2-F1 non-sig	T1-T2- F2 non-sig
			T2-T3-F1 $p < .001$	T2-T3-F2 $p < .001$
			T1-T3-F1 $p < .001$	T1-T3- F2 $p < .001$
/u:/	$p < .001$	$p < .001$	non-sig	T1-T2- F2 $p < .001$
				T2-T3-F2 $p < .001$
				T1-T3- F2 $p < .001$
/e/	$p < .001$	$p < .001$	T1-T2-F1 non-sig	T1-T2-F2 $p < .015$
			T2-T3-F1 $p < .001$	T2-T3-F2 $p < .001$
			T1-T3-F1 non-sig	T1-T3-F2 $p < .001$

Overall, the above results indicate that the vowel /u:/ for F1 did not undergo significant changes over time; however, the vowel's F2 did show important variations over the three tests, with a tendency toward a decreased F2 values at T2, but higher F2 values at T3, the latter being significantly higher compared to that produced in the initial test, T1. The results for the vowel /ʊ/ show a significant trend toward an increased F1 across T2 and T3; a lower F2 by T2, but significantly higher F2 values by T3.

For the vowel /i:/, F1 and F2 did not undergo statistically significant changes over time. The vowel /ɪ/, however, shows a substantial increase of F1 values by T2, and a significant decrease in F1 by T3, returning to a value similar to that produced at T1. Concerning F2, the segment was produced with a lower F2 by T2, but a significant increase in the value by T3,

almost at the same level as for T1. The vowel /e/ showed significant changes in F1 at T3 and increased F2 at T3.

The next section provides the results of vowel length for /i:/ and /ɪ/, /u:/ and /ʊ/.

### 5.5.2 Duration results

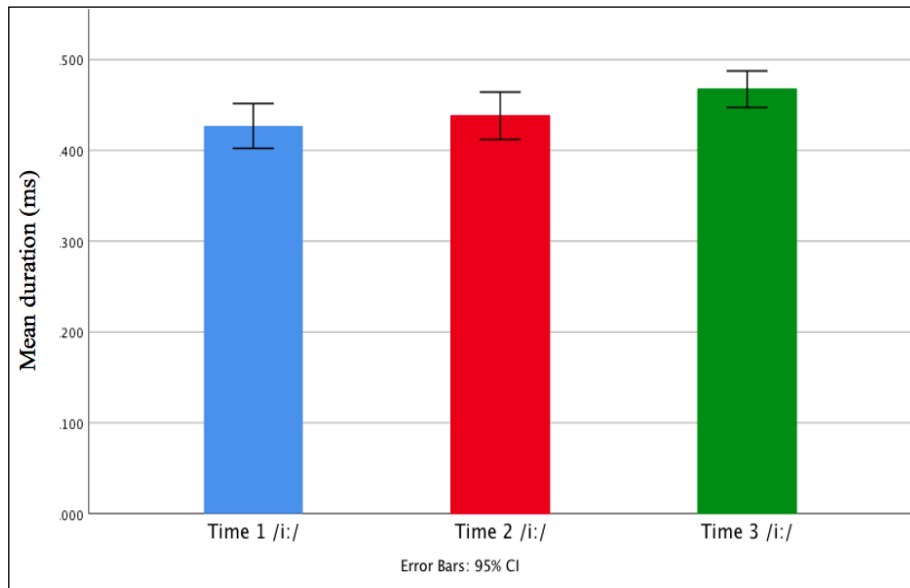
The mean duration of the four vowels over the three time points was calculated, and then a Repeated Measures ANOVA was conducted. The within-subject effects for the ANOVA were time (3), vowel (4) and vowel \* time. Interactions between vowel and time, and distance between vowel contrast and time, were followed by post-hoc analyses: pair-sample t-tests ( $p < .05$ ) for which all significant results were Bonferroni corrected.

The ANOVA results show that vowel lengths and vowel-time interaction significantly differed over the time points: time:  $[F(2, 78) = 19.271, p < .001]$ , vowel:  $[F(3, 117) = 4.352, p < .006]$ , and vowel-time interaction:  $[F(4.654, 181.498) = 42.763, p < .001]$ . Post-hoc results for the vowel /i:/ over time indicated a significant duration difference between T1 and T3,  $p < .036$ ; on the contrary, results for T1-T2 and T2-T3 did not reach the level of significance (see Figure 14). Regarding the vowel /ɪ/ the duration over the time points revealed significant differences between T1-T2, and T2-T3 at  $p < .001$ ; whereas the difference found between T1-T3 did not reach statistical significance (see Figure 15). Finally, the post-hoc analyses for the distance between the pair of contrast vowels /i:/-/ɪ/ over the three time points showed a significant effect in T1, T2 and T3, all  $p < .001$  (see Figure 16).

**a) Vowels /i:/-/ɪ/**

**Figure 14**

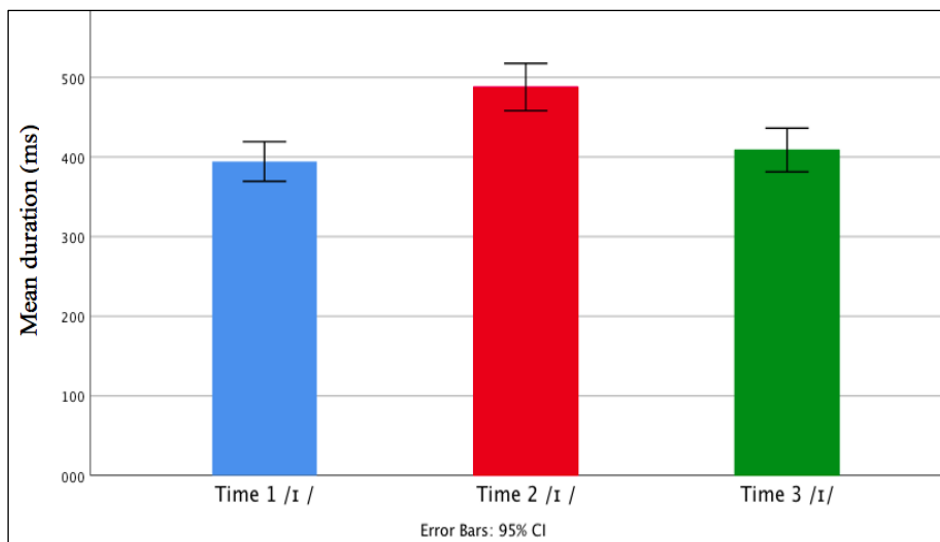
*Mean duration of vowel /i:/ for Time 1, 2, and 3*



Note: Mean across times: 444 ms

**Figure 15**

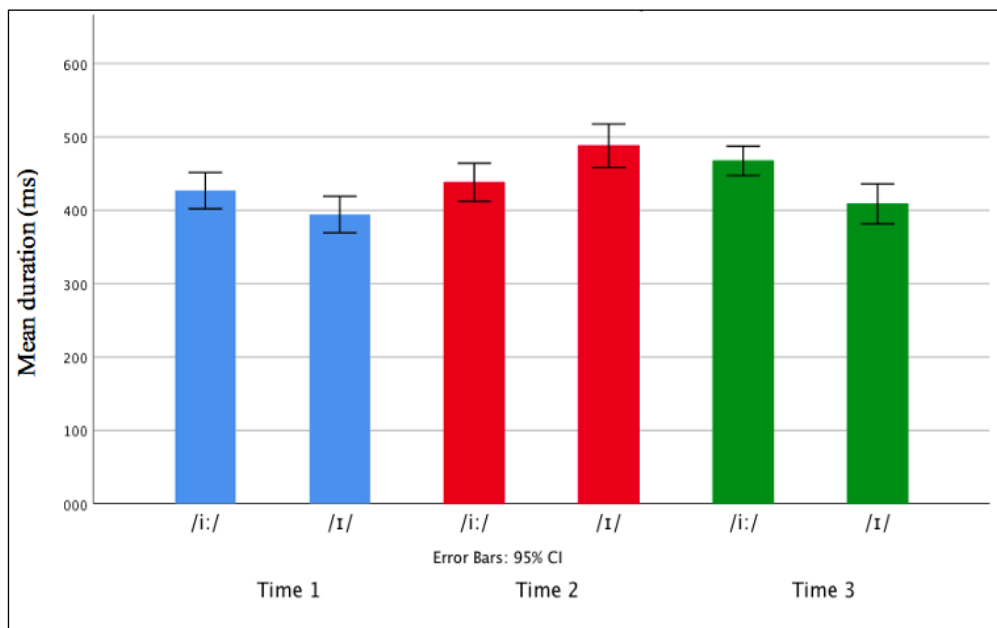
*Mean duration of vowel /ɪ/ for Time 1, 2, and 3*



Note: Mean across times 432 ms

**Figure 16**

*Vowel contrast duration /i:/-/ɪ/ for Time 1, 2, and 3*

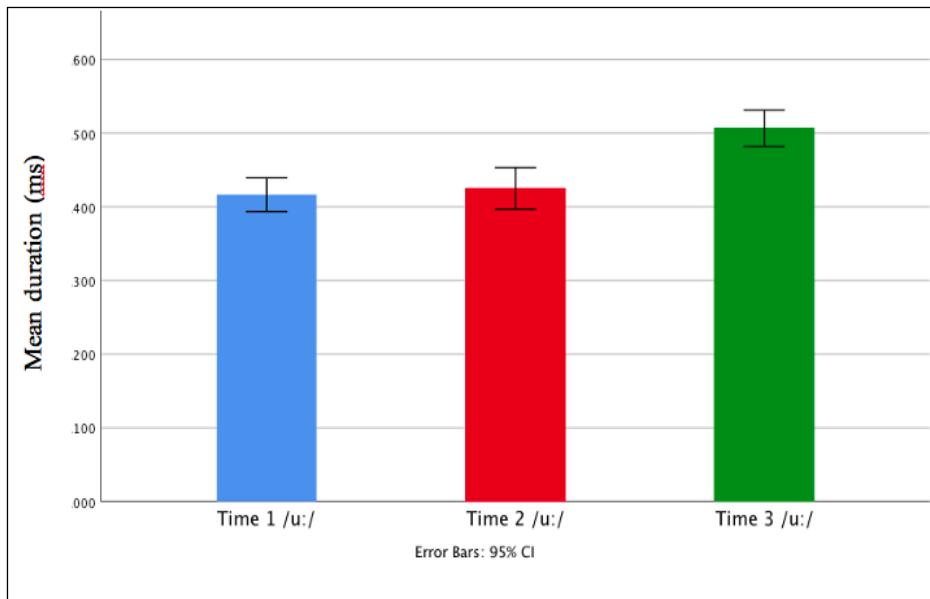


**b) Vowels /u:/-/ʊ/**

In addition, post-hoc analyses for the vowel /u:/ between T1 and T2 showed no significant differences (see Figure 17). The results between T2 and T3, and between T1 and T3 showed significant differences at  $p < .001$ . Post-hoc tests for the vowel /ʊ/ showed significant differences for T1-T2, T2-T3, and T1-T3 interactions at  $p < .001$  (see Figure 18). Moreover, the distance between the pair contrast /u:/-/ʊ/ and time points showed significant results only at T1 [ $p < .036$ ], while the outcomes for T2 and T3 did not reach statistical significance (see Figure 19).

**Figure 17**

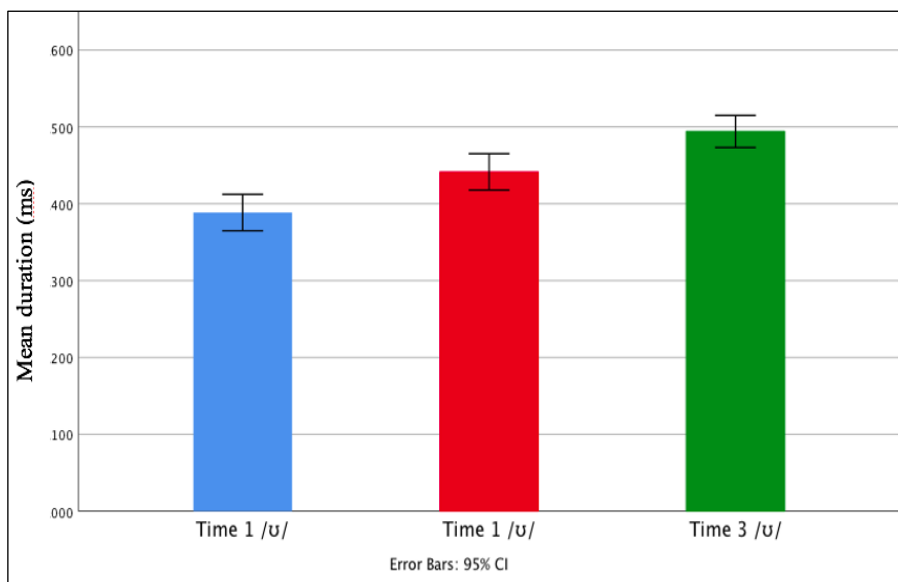
*Mean duration of vowel /u:/ for Time 1, 2, and 3*



Note: Mean across times 449ms

**Figure 18**

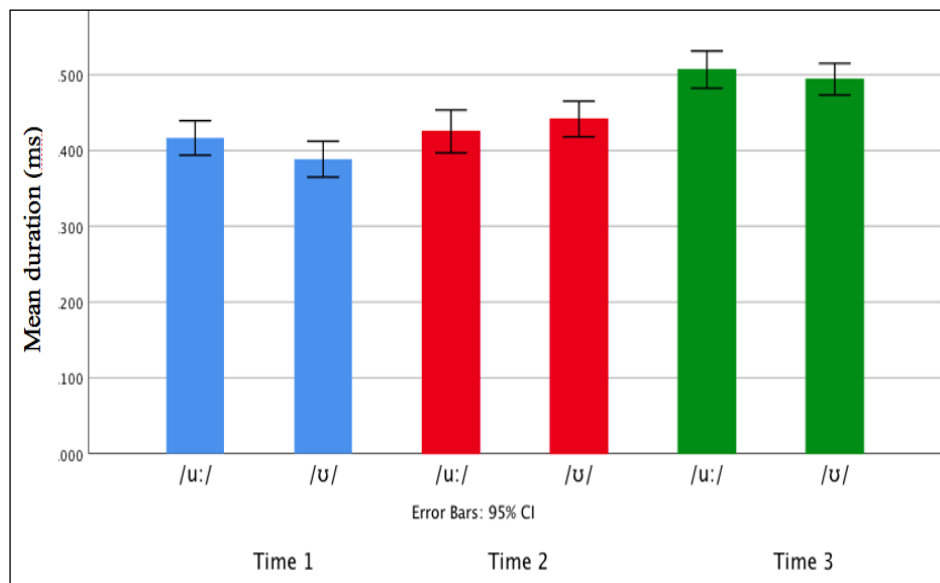
*Mean duration of vowel /ʊ/ for Time 1, 2, and 3*



Note: Mean across time 441 ms

**Figure 19**

*Vowel contrast duration /u:/-/ʊ/ for Time 1, 2, and 3*



The following Table displays a summary of the results.

**Table 9**

*Summary of the Post-hoc test results*

Post-hoc	Times		
/u:/	T1-T2 non-sig	T2-T3 $p < .001$	T1-T3 $p < .001$
/ʊ/	T1-T2 $p < .001$	T2-T3 $p < .001$	T1-T3 $p < .001$
/u:/-/ʊ/	T1 $p < .036$	T2 non-sig	T3 non-sig
/i:/	T1-T2 non-sig	T2-T3 non-sig	T1-T3 $p < .036$
/ɪ/	T1-T2 $p < .001$	T2-T3 $p < .001$	T1-T3 non-sig
/i:/-/ɪ/	T1-T2 $p < .001$	T2-T3 $p < .001$	T1-T3 $p < .001$

The above results suggest that for the length of /i:/ there was a non-significant increase between pairs of time points (T1-T2, T2-T3). However, overall, the duration of the produced vowel increased significantly over the year (from T1 to T3). This suggests a strong time effect



for the segment. For /ɪ/ the outcomes suggest an appreciable increment in vowel length by T2, followed by a significant reduction of the segment duration by T3, almost reaching the same length measured in T1. The distance between the pair of contrastive vowels reveals significant results over time, suggesting that time affected the production of the segments. Interesting results were found at T2, where the realisation of /ɪ/ was longer than /i:/, following the same pattern as the /u:/- /ʊ/ results. Overall, by T3 the distance between the /i:/ and /ɪ/ vowels was larger than at the previous two testing times.

The length of /u:/ increased across the three-time points examined. Although the difference between T1 and T2 was not significant, the production of the vowel was nonetheless longer; the comparison of T2 and T3 demonstrated a significant increase, showing the effect of time on the production of this segment. The results for the vowel /ʊ/ reveal a significant length increase at both testing times. Interestingly, the outcomes on the vowel pair contrast showed that at T1 the durational difference between /u:/and /ʊ/ was significant; however, by T2 and T3 this distance was non-significant and /ʊ/ was produced longer than /u:/ during T2. Overall, both vowels show a clear tendency towards a length increment over time (see Figure 19).

## 5.6 Summary

The aim of the present chapter was to examine the production of the English vowel contrasts /i:/ and /ɪ/, /ɪ/ and /e/ and /u:/ and /ʊ/ with respect to both spectral and temporal features by adult Spanish speakers (as a group) in a citation style of speech, to obtain an overview of the general performance of the group with regard to the British English targets before looking at individual performance.

Overall, in terms of formants by T3, the group - as a whole - obtained results above the reference values from the contemporary RP data . However, the general distance between the pair of vowels /i:/ and /ɪ/ and /u:/ and /ʊ/ produced by the Spanish speakers was below the English targets; and the distance between /ɪ/ and /e/ was greater than the English values. (see Appendix E)

### 5.6.1 Formant analysis (vowel quality)

To summarise the group performance at the beginning and end of the year as measured by F1 - F2 Euclidean distances:

*/i:/ vs /ɪ/*

- At T1, the group did not have a clear distinction between this pair at the beginning of the year (E. distance = 112 Hz) and had not developed one by the end (E. distance = 235 Hz).

*/ɪ/ vs /e/*

- At T1, the group had a clear distinction between this pair at the start of the year (E. distance = 268 Hz) and had maintained their clear distinction at the end of the year (E. distance = 394 Hz).

*/u:/ vs /ʊ/*

- At T1, the group did not have a clear distinction between this pair at the beginning of the year (E. distance = 68 Hz) and had not developed one by the end (93 Hz).

To summarise the group performance at the beginning and end of the year as measured by F1 and F2 separately.

- By T3, the vowels */ɪ/* and */ʊ/* showed significant differences in their F1, F2 structure.
- By T3, */e/* significantly increased its F1 and F2.
- By T3 */u:/* significantly increased its F2.

### 5.6.2 Vowel duration

- Results (T1-T3) between */u:/* and */ʊ/* show that the duration was not significant.
- Results (T1-T3) between */i:/* and */ɪ/* show that the difference in duration was statistically significant.
- Although some results are statistically significant, the group did not make a vowel quantity distinction. Some speakers, in fact, produced the short members of these pairs

fractionally longer than the long members; others made the distinction in the direction associated with native English, but the durational differences in all cases only involved only 10 ms.

**Table 10**

*Vowel duration values*

	<b>English</b>	<b>Group results (Spanish speakers)</b>
Short vowels	160 ms	436 ms
Long vowels	300 ms	446 ms

### **5.6.3 General**

This chapter has provided a general view of the group performance for the production of the English vowels at three points in time. A discussion of the findings set out in this chapter is deferred until the final Discussion chapter.

The following Chapters examine differential degrees of progress towards a native-like pattern in vowel production across time at the individual level. These are related to factors that may have impeded or facilitated adaptations towards native English productions.

## **Chapter 6. Marking of English Vowel Contrasts over a Year: Individual performance**

### **6.1. Introduction**

The present chapter builds on chapter 5 to determine the degree of separation that individual speakers obtained across time points between contrasts in the English vowel pairs /i:/ and /ɪ/, /ɪ/ and /e/ and /u:/ and /ʊ/ with respect to both spectral and temporal features using citation-style speech. The individual vowel distance achieved by the speakers is presented in terms of Euclidean distances (distances between mean points on the F1/F2 plane).

For convenience to the reader, Tables and Figures to illustrate the different trajectories for individual speakers over time appear in Appendices F and G.

#### **6.1.1 Structure of the chapter**

This chapter is organised into the following sections<sup>4</sup>:

- contextual information
- vowel distance analysis
- individual progress/distribution results
- individual formant results
- /u:/ and /ʊ/ fronting results
- vowel duration outcomes
- summary

---

<sup>4</sup> Method: stimuli, recording procedures and participants are the same as those presented in Chapter 4.

## 6.2 Elements affecting L2 vocalic perception and production

### a) Native English speakers

Cue weighting relates to the relative importance of acoustic properties that speakers of a language pay attention to when they perceive contrasting sounds.

For native English speakers the key, i.e., most heavily weighted, cues are spectral and durational, while others, although present, play a secondary role at the moment of segment identification (Holt & Lotto, 2006). Most research agrees that when native English speakers discern vowel contrast, particularly between the high vowels, they tend to focus on spectral cues over temporal cues (Hillenbrand et al., 2000). A study conducted by Flege et al. (1997) regarding identification responses to vowels found that native English speakers made significantly more use of spectral quality than temporal cues to detect vowels, stating that durational cues were used primarily when spectral signals were insufficient or vague. Moreover, Hillenbrand et al. (2000) investigated the role of duration in vowel perception. 12 vowels in a /hVd/ context were included in four sets of synthesised words with different vowel durations. Fifteen phonetically trained listeners participated as subjects. The results suggested that for vowels which present consistent durational differences but an adequate degree of spectral separation, such as /i:/ and /ɪ/, /u:/ and /ʊ/ or /ɪ/ and /e/, duration plays a minimal role in perceiving vowel differences. However, for vowels that are closer in their spectral features, such as /ɑ/-/ɔ/-/ʌ/ or /æ/-/ɛ/, durational cues are more important for the listeners. Therefore, the greater the formant separation between vowels the lower the reliance on durational cues to identify segments, and vice versa. A study conducted by Zahorian & Jagharghi (1993) concluded that in vowel recognition, listeners, regardless of the separation between segments, depend more on the quality of vowels rather than the temporal features, the latter being used only as a secondary source of information. In addition to the previous arguments, Escudero & Boersma (2004) pointed out that speakers of different English dialects might pay attention to different acoustic cues. In their research, they tested the identification of the vowel contrasts /i:/ and /ɪ/ by L1 Scottish and Southern British English (SBE) speakers. The results showed that for the group of Scottish English speakers, spectral cues were more significant than durational ones, whereas the Southern British participants showed equal dependency on both temporal and spectral cues for the correct distinction of the high front pair of vowels. The difference in the group's results is explained in terms of dialects; since

Scottish English speakers produce a length distinction between the vowel pair /i:/ and /ɪ/ depending on the voicing of the following consonant, and whether there is a following morpheme boundary. In the identification task speakers relied mainly on spectral cues disregarding duration, an action that differed from SBE speakers due to their /i:/ and /ɪ/ phonological features. Overall, researchers have identified and mostly agree that native English speakers predominantly use spectral signals to identify vowels, placing vocalic duration as a superfluous or secondary feature for vowel differentiation.

### **b) Non-native English speakers**

In second language learning, specifically in late learners, the phonology of the L1 interferes with how the L2 sounds are perceived; therefore, difficulties in L2 speech comprehension are linked to the lack of resemblance of the phonological categories between the L1 and the L2. An example is the case of Japanese speakers of English and their difficulty in perceiving and producing the English phonemes /l/ and /r/, a problem that has been attributed to the assimilation of two different L2 categories into a single L1 class (Hattori & Iverson, 2009; Yazawa et al., 2020). It is important to note that regarding adult second language learning, attention has been given to speakers of languages that, in contrast with English, lack a tense-lax vowel system, such as Mandarin or Spanish, to determine how they either achieve or fail to produce vowel contrasts in their L2. According to Kim et al. (2018), learners of English as well as native English speakers use acoustic cue-weighting to identify vowel differences, favouring spectral or durational cues. The selection and learning of cue weights are essential for the development of speech comprehension, phonological categories and later speech production in the L2 (Nittrouer & Lowenstein, 2010). However, speakers of languages where temporal vowel differentiation do not play a principal role in contrasting segments might face different challenges when accessing or understanding acoustic signals, since they might depend on different cues from those used by native English speakers. For example, in the case of native Spanish speakers who, at least in their initial stage of L2 learning, rely on vowel duration rather than spectral cues to perceive the English tense/lax vowel contrast, while native English speakers do the opposite (Escudero, 2000; Kim et al., 2018; Kondaurova & Francis, 2010).

The desensitization hypothesis developed by Bohn (1995) accounts specifically for speakers of languages which, unlike English, might not be sensitive to L2 spectral distinctions, for example in high front vowels /i:/ and /ɪ/, thus establishing that if listeners of an L2 do not have experience of formant differences, they will become *desensitized* towards them, and will use mainly temporal signals in their production to show vowel contrast, as in the case of L1 Spanish speakers. Escudero & Boersma (2004) argue that the fact that Spanish learners of English pay attention to duration rather than spectral cues in tense/lax vowel contrasts is simply because their L1 does not have duration categories; therefore, the use of temporal cues to classify vowels in another language is open while spectral classifications are not. This argument can be supported by the findings from Escudero et al. (2009), which state that Spanish speakers learning Dutch rely on duration rather than the spectrum for the /a:/-/ɑ/ vowel distinction, a phenomenon that shifts as the L2 language level improves. Additionally, Escudero (2000) suggests four stages for the development of the L2 perceptual weighting structure: (a) initially, speakers will have no perception of the contrast between vowels such as /i:/-/ɪ/; (b) next, the contrast will be perceived but only in length; (c) later on, both spectral cues and segment duration will be taken into account, with duration as the primary signal; (d) and finally, speakers will manage to give the appropriate weight to durational and spectral cues in a way that is similar to what native speakers do. It could be argued that these stages develop as exposure and contact with the L2 increases and cue weightings come to be balanced consistently with the target language. In support of this, Escudero & Boersma (2004) found that Spanish speakers exposed to different English dialects (Scottish vs Southern British) used different acoustic cues to perceive the English vowel contrast /i:/-/ɪ/. Spanish participants exposed to Scottish English as their target language relied on spectral cues as native Scottish English speakers do; however, participants who were exposed to Southern British dialect unlike native speakers, who used both spectral and durational cues, tended to focus mainly on duration. This divergence is explained by the Spanish speakers' reliance on duration alone, despite their proficiency in the L2.

In addition, Wang & Munro (1999) tested the perception and production of the English vowel contrast /i:/-/ɪ/ and /u:/-/ʊ/ in 14 adult Mandarin speakers (graduate students living in Canada). The participants underwent a pre-test, cue weighting training, and post-training test. The perception results for the vowel contrasts showed that during the pre-test participants relied mainly on temporal cues. However, after a few training sessions, the subjects improved their use of spectral signals, especially for the front vowels. The outcomes for the production of the

vowels indicated an over-reliance on duration on the /i:/-/ɪ/ vowel contrast, yet this inflated reliance on duration was not applied to the back pair /u:/-/ʊ/. Inconsistencies in the use of temporal cues were noted, but there were also beneficial results in awareness of vowel quality signals, although the long-term benefits of acoustic cue training are yet to be proven and explored.

In general, internal and external factors play a role in the development of new phonological categories in a second language learning experience. There are several interconnecting factors which are formed at the moment of producing and perceiving new sounds such as the learner's L1, the exposure (input) to the target language, the length of residency (in the country where the L2 is spoken) and age, among others. This section reviews different points of view regarding these features.

### **6.2.1 Exposure, quality and quantity of L2 input**

In general, internal and external factors play a role in the development of new phonological categories in a second language learning experience. There are several interconnecting factors which are formed at the moment of producing and perceiving new sounds such as the learner's L1, the exposure (input) to the target language, the length of residency (in the country where the L2 is spoken) and age among others. This section reviews different points of view regarding these features.

### **6.2.2 Input**

The role that length of exposure plays in L2 English learning is not a settled argument. L2 input has been mostly defined as utterances that have been received and understood by the L2 learner irrespective of the origin. Therefore, if the speech is comprehensible for the listener, it can be produced by native or non-native speakers of the target language in different settings, such as inside a classroom or in a natural environment outside the classroom (Flege, 2009; Leow, 2007).

Some researchers claim that the input provided by native speakers to a L2 English learner plays a restricted role, stating that exposure to the target language (especially after puberty) does not have a strong effect on how the new segments are produced, perceived or modified



by the L2 learner. From this perspective, the learner's age, at the moment of first exposure to a new language, is placed as the main responsible factor for L2 outcomes (DeKeyser, 2000). In contrast, the idea that the quality and quantity of input play a central role in segment modification has been formulated by other scholars. According to Flege (2018), changes in contrasting English vowels are possible for early (young) and late (adult) L2 learners as long as they have substantial exposure to the target language. Flege argues that exposure alone is not enough for the L2 speaker, but the quantity and quality of the input play a key role in bolstering L2 learning, even beyond the effects linked to the age of acquisition.

L2 exposure happens principally in two environments: the formal classroom, with a limited time of language use, and in a natural setting, by total immersion, where the learner can receive constant input and use the language frequently (Best & Tyler, 2007). This constant input for non-native speakers can differ in quality and quantity. According to Flege (2009), formal classroom instructions, that usually happen in the learner's home country, often present a lack of quality in the amount of input offered to the students. This observation is based on the fact that English teachers are usually non-native English speakers.. On the other hand, learners in total immersion settings such as migrants arriving in English-speaking countries are exposed to highly diverse L2 input in terms of both quality and quantity; however, this language influence is conditioned to the L2 speaker's social circle and the necessity to interact in English. Therefore, the quality of the language and what L2 learners are exposed to can vary, because the language exchange can come from different sources, such as non-native speakers, compatriots or native English speakers, who have different backgrounds and dialects.

In support of this, Flege & Wayland (2019) investigated the effect that different input types had on the production and perception of English phonemes among non-English speakers. The participants were 60 native Spanish speakers who arrived in the United States after puberty (16 years of age). The subjects were divided into low, medium, and high input groups according to their length of residency in the US and the English input they declared to have received. The results regarding vowel perception showed similarities across the different groups and, in terms of vowel production, the high input group recorded a modest improvement in comparison to the other two groups. The outcomes were explained based on two possible interpretations: first, the possibility that the input received has a limited effect when the L2 is learned in a naturalistic context, but after the critical period; and second, adequate native speaker input is necessary to support L2 learning. A criticism of the second

interpretation was that the input was not sufficiently evaluated in the research. By contrast, Jun & Cowie (2004) tested adult Koreans, with different times of residency in the US, to assess if time of exposure and input would benefit their English vowel production. Results showed that experienced Korean speakers (26-31 years living in the US) outperformed the less experienced speakers (1-5 years in the US), supporting the argument that sufficient quantity, as well as good quality of English language, helps the development of L2 segment production.

### **6.2.3 Second language experience**

Second-language experience, although a general term, has principally been defined as the multiple factors that can influence second language performance (Gorba & Cebrian, 2021). Empirical studies have suggested that L2 learners' perception and production of non-native speech vary according to their L1; however, these in turn have been shown to develop as the learners' experience with the target language increases (Best & Strange, 1992; Flege et al., 1997). In support of this, Lee & Cho (2000) studied the effect that L2 learning experience has on non-native speakers' perception of English vowels. In their study, Korean participants living in the US were divided into two groups according to the length of residence (4 vs 11 years). The results showed that overall, the group with a longer time of residence in the US outperformed, in terms of identifying English segments, the less experienced group. These outcomes support the view that involvement and familiarity with the target language is crucial in vowel perception tasks. In addition, Flege et al. (1997) studied a group of German, Mandarin, Korean, and Spanish speakers to evaluate the production and perception of /i:/, /ɪ/, /e/ and /æ/ vowels. The groups were divided into experienced and inexperienced speakers, according to their length of residence in the USA (9 years vs 4 months respectively). The results indicated that experienced learners produced and perceived the English vowels under examination more precisely than the naïve learners. In addition, the specific findings for production indicated that inexperienced Korean, Mandarin and Spanish speakers did not produce spectral differences for the vowel /i:/- /ɪ/ contrasts while Korean, Mandarin and German speakers did not use spectral differences for /e/ and /æ/. In terms of perception, the inexperienced participants mainly relied on the duration for /i:/- /ɪ/. The findings support Bohn's (1995) idea that duration is primarily used when there is no spectral equivalent in the

speaker's L1. In addition, this study showed that, for the experienced learners, increased exposure to the L2 had an effect on vowel production and perception. Moreover, it could be argued that, despite speakers' L1 vowel inventories, they all improved in their use of temporal and spectral signals as they gained experience with the target language.

In contrast to the previous arguments, a study by Cebrian (2006) proposes a disassociation between the importance of exposure (experience) and improved performance in an L2. Cebrian tested two groups of Catalán learners of English (group 1: 25+ years in Canada and group 2: L2 learners living in Barcelona, Spain) to trace how temporal and spectral cues, which are used to distinguish vowel /i:/-/ɪ/ contrasts, develop. The results showed that, although the two groups greatly differed in exposure to the target language, they both mainly depended on durational cues and failed to use spectral signals for vowel identification. These findings support the desensitization hypothesis and do not show a correlation between L2 experience and the development of spectral differences to distinguish vowel contrasts. In addition, Grenon et al.'s (2019) study of 23 adult Japanese learners of English who underwent explicit /i:/-/ɪ/ vowel contrast training for two to five weeks (in Japan), found that half of the participants achieved vowel distinction after the training. The authors posited that awareness and directed phonetic training in spectral cues might produce better results than increased experience via immersion in the English language, at least for adult learners. Casillas (2020) tested English learners of Spanish over a 7-week training period, with high exposure to Spanish and no use of English. The results show that English speakers improved their Spanish stop consonants showing that, at the initial stages of learning, the constant use of the L2 has a favourable impact on the phonetic learning of production and perception. Importantly, these studies, although successful, fail to support the idea that participants will maintain or continue to improve their perception and production of the L2 targeted segments after the sessions cease.

#### 6.2.4 Length of exposure to the target language

Total immersion in the L2 country is an essential factor for language performance, especially for adult learners. In addition, length of exposure, by means of residency in the L2 country, has been positively associated with better English language performance (Guion et al., 2000). The exact time or minimum length of residency (LOR) to accurately develop foreign segments is still a debatable topic, since research findings have been mixed and have not reached a consensus on the topic.

LOR is not automatically linked to good quality and quantity of English input; therefore, developing a new language is not just a matter of being immersed (length of exposure) in an L2 environment. For example, learners with an equivalent residence time but with a different use of L1 and L2 may present variations in their linguistic development (see Piske et al., 2001).

Regarding the length of residence, Baptista (2006) evaluated the production of English vowels by 11 Brazilian-Portuguese speakers residing for one to six months in the US. The results of Baptista's monthly recordings indicated that 9 out of 11 participants were not able to produce /i:/-/ɪ/ vowel contrasts after four to six months of exposure. However, few participants managed to improve formant frequency of other vowels tested after eight months. These results indicate that eight months of exposure were insufficient for vowel contrast development for the majority of the speakers tested. Similarly, Song & Eckman (2019) measured the production and perception of the English vowel contrasts /i:/-/ɪ/ from adult speakers of Spanish, Portuguese, and Korean (36 participants in total). They participated in three sessions of tests every four weeks in addition to attending training sessions to enhance awareness of the vowel contrasts examined. The results revealed that half of the participants showed vowel distinction; the specific results of the Spanish speakers, found that three out of seven subjects produced the vowel contrasts /i:/-/ɪ/ throughout the three sessions, and six participants show contrast only in the formant values in at least in one session; these results indicate that participants were on an intermediate path to producing the correct phonemic difference. In terms of perception, two subjects were able to perceive the distinction at the end of session three. It should, however, be noted that in this instance attention was paid to durational cues over spectral ones. The main findings suggested that in order to produce this vowel contrast, three months of exposure is not sufficient, and a further period of time is

needed to have a sustained performance. The results of the studies conducted by Morrison (2002), which tested the development of vowel contrasts /i:/-/ɪ/ by Japanese and Spanish speakers living in Canada, support the view that less than five months of exposure is not sufficient time to develop awareness of spectral segment differentiation.

Further support for increased length of exposure can be found in the one-year longitudinal study by Munro & Derwing (2008) whose main objective was to examine the development of English vowel production. The participants were 44 adults (20 Mandarin speakers and 24 speakers of a Slavic language) who were living in Canada for fewer than four months at the time of the study. Vowel production was measured by trained and untrained listeners (native English speakers) at two-month intervals. The results indicated that participants improved their intelligibility in vowels, especially within the first six months, supporting the proposal that adult learners can improve performance in segment acquisition with time (Flege et al., 1992). The overall results for some speakers were aligned with the suggestion of a rapid progress of L2 segments at an initial stage of immersion followed by a plateau effect (Flege, 1988) suggesting that, although improvements in the intelligibility of certain vowels were seen at the end of the year, the progress across the vowels examined was not homogeneous for the speakers. Interestingly, the results from this study, contrary to those obtained from Baptista (2006) and Song & Eckman (2019), show a promising framework of adequate length of exposure for L2 segment development. The subjects were tested for a longer period of time allowing one to reach the conclusion that the first six months of learning a second language (over a one-year period) are crucial for the development of L2 vowels.

In contrast, a study by Baker & Trofimovich (2006) which tested Korean speakers (in university), who had been living in the US with varying lengths of residence (less than 1 year, 3 years and 10 years), found that the improvement in English vowel production mainly happened among those in contact with the L2 for more than 10 years. Less than 3 years residing in the US was found to be not enough time for sufficient changes in segment accuracy. In a similar vein, Smith et al (2019) compared the formant values produced by Mandarin, Korean, and Spanish speakers, living in the US for 2 to 12 years, to those of native speakers of American English. The general results indicated that non-native speakers differ significantly from native English ones in F1 values but not in F2 values within the tokens examined. The authors therefore suggested that, despite the length of residence and exposure to the L2, the outcomes might reflect fossilisation of the segments targeted. A longitudinal case study by Koffi & Lesniak (2019) tested a Spanish speaker producing English vowels, the

results showing a clear adaptation towards producing different English vowels for the front and central ones; however, the back pair /u:/ and /ʊ/ did not present a spectral distinction, and both vowels had an overlapping production. The authors suggested that back vowels undergo a slow process of modification.

As these findings reveal, length of residence does not necessarily correlate with changes in all L2 segmental production, and it might confirm the claim that, beyond a certain time of exposure, pronunciation changes cease to occur (Piske et al., 2001).

### **6.2.5 Native language (L1)**

Research has shown that a L1 phonetic background, especially for adult learners, can influence the speed or the degree of accuracy with which non-native English speakers produce and perceive English tokens (Flege, 1995). The influence of the L1 phonological system over the L2 differs and mostly depends on the native language and the type of L2 phonological contrasts that the learner is perceiving or producing (Polka, 1991).

In the case of Spanish, the vowel inventory of the language presents less diphthongality when compared to English and does not show duration of vowel contrasts (Flege et al., 1997).

According to Escudero & Chládková (2010), some assimilation problems might be present when contrasting Spanish and English vowels, especially during the early phases of learning.

The results obtained in their research propose that a Spanish speaker's identification of English F1- F2 values will resemble the Spanish counterpart, and in turn, they suggested the following perception pattern (above 30-70% assimilation) for L2 learners of British English:

**Figure 20**

*Possible Assimilation patterns for British English produced by Spanish speakers*

SSBE		Spanish
/i:/	→	/i/
/ɪ/	→	/e/
/ɛ/	→	/e/
/æ/	→	/a/
/ɑ/	→	/a/
	↘	/o/
/ɔ/	→	/o/
/ʌ/	→	/a/
/ʊ/	→	/u/
/u:/	↗	

(Adapted from Escudero & Chládková, 2010)

The above prediction (Figure 20) can be exemplified for vowel contrast in English words such as *feel* - *fill*, where Spanish speakers would likely perceive or produce *fill* as the past tense of the verb *fall*, *fell*, reflecting the Spanish /e/ as in *él*. Another vowel distinction prone to assimilation could occur in the minimal pair *pool* – *pull*, leading to a production or perception of these different sounds as a Spanish /u/, such as in the pronoun *tú*. Also, the vowel /ɑ/ as in *calm* would probably be perceived and produced as the Spanish /a/ as in the word *pan* or as in the vowel /o/ as in *sol*. Previous acoustic analysis studies, such as Flege (1991), have demonstrated a similar pattern as those shown above. This is particularly true for the English vowel /ɪ/ produced as Spanish /i/, /ɛ/ realised as the Spanish /e/, and the English vowel /æ/ assimilated to the Spanish /a/. Furthermore, Escudero & Chládková (2010) suggest that Spanish speakers producing British English face problems with the back vowel contrast /u:/- /ʊ/ due to the need to divide a single L1 category (/u/) into two, in addition to having to modify acoustic signals to achieve higher F2 values.

A study by Fullana-Rivera & Mackay (2003) examined the production of the English vowel contrast /i:/ and /ɪ/ by a group of Spanish speakers (N=135) and tested their performance by

means of an identification test performed by native English speakers. The participants recorded a list of 34 English words containing the vowels examined. The results showed that most of the Spanish speakers produced a distinctive /i:/ with native-like frequencies; however, the /ɪ/ was found to be mispronounced with frequencies similar to the /i:/. The results from the identification task (by native English speakers) show that the production of /i:/ was most of the time correctly identified with a range between 58% -100%. By contrast, the correct identification of /ɪ/ was low, with positive responses between 38% - 60% range, and it was most of the time confused with /i:/. These findings agree with previous research e.g., Flege (1991) and Flege et al. (1997), that suggests that Spanish speakers produce /ɪ/ with values closer to /i:/.

Overall, there seems to be evidence to indicate that phonetic interference stemming from different language backgrounds is an important element in L2 segment production.

Together, these studies have reported inconsistencies related to the importance of input and the amount of time needed for non-native English speakers to show vowel contrast differentiation. Most of the studies support the notion that adult learners are capable of modifying the production of segments, after considering individual differences such as motivation, age, social interaction, and L1 background, which play a role in L2 speech development.

### **6.3 /u:/and /ʊ/ vowel fronting in British English**

Over the last fifty years, high back vowels /u:/ and /ʊ/ have shifted in many English accents to a more fronted position, including the Standard accent of England. This process has affected both /u:/ and /ʊ/ vowels in non-linear time, with /ʊ/ fronting being a more recent phenomenon. The fronting of these vowels is not always social-class indexed in places where the advancement occurs. Age is an important factor in the realisation of the tokens, where younger speakers tend to produce more fronted vowels compared to older speakers (especially in SSBE or RP) (see Harrington et al., 2008; Hawkins & Midgley, 2005; Strycharczuk & Scobbie, 2017). The phonetic context plays a key role in the fronting of the two vowels; coarticulation with preceding coronal or palatal consonants promotes fronting, while a following coda /l/ reduces or obstructs it (Kleber, et al., 2011; Sóskuthy et al., 2015). In addition, tongue movement has been generally seen as the main indicator for the rise of the F2 values; however, lip unrounding could also explain the change in articulation. Hughes et al.



(2012) describe the RP /u:/ lip rounding as a variable aspect among speakers, while /ʊ/ lip-rounding is neutral or close. Fabricius (2007), however, describes /ʊ/ as a totally unrounded vowel, when produced by young speakers. Studies indicate that age is a primary factor in the ongoing change of acoustic quality of /u:/ and /ʊ/. To compare formant values, Hawkins & Midgley (2005) examined 20 RP speakers who had been divided into four age groups (20-25 yr., 35-40 yr., 50-55 yr. and 60-73 yr.). The results showed an age-related increment in the F2 values for both /u:/ and /ʊ/. The pace of the increase, however, was dissimilar between them. Whilst speakers of all age groups produced /u:/ with higher F2 values, only the younger group (20-25 yr.) produced /ʊ/ with noticeably higher F2 values relative to the oldest's group productions. These results were supported by a study conducted by Harrington et al. (2008), which established the fronting of the /u:/ in Standard Southern British speakers (SSBE). In this study, participants were grouped according to their age: *the young group* (18-20 yr.) and *the old group* (50+ yr.). The results for the production of /u:/ showed that, although both groups fronted their vowels, the young group had an increment in the F2 values almost at the same level of the vowel /i:/, a tendency that was not observed in the older group.

Table 11 shows a comparison of F2 values across different studies; it indicates that speakers at the moment of producing /u:/ and /ʊ/ have undoubtedly increased their second formant over past decades.

**Table 11***Comparative table of F2 frequencies from different studies*

Source	Speaker's age	Avg. F2 /u:/	Avg. F2 /ʊ/	Avg. F2 /i:/
Wells (1962)	18+	939	950	2373
Deterding (1997)	Unspecified	F:1437	F:1340	F: 2554
		M: 1191	M:1550	M: 2249
Ferragne & Pellegrino (2010)	18-30	F: 2202	F: 1705	F: 2760
		M:1672	M: 1550	M: 2289
Bjelaković (2016)	50 +	F: 1853	F: 1492	
		M: 1684	M: 1345	

Note: M= male, F= female subjects.  
 (Adapted from Strycharczuk and Scobbie, 2017)

## 6.4 Euclidean distance

Euclidean distance measurements have been used in linguistic research to compare vowel systems and measure vowel distances and changes in L2 pronunciation (Mairano et al., 2020). This method calculates, in a simple way, the distance between two vowels by using the means of the F1 and F2 values for each category. The results, which can be in any unit of measurement, show how distant two vowels are produced by a speaker, with easy visualisation by a line between dual points (Nycz & Hall-Lew, 2013).

The formula below was used in the following analyses and in the ones presented in Chapter 8:

$$d = \sqrt{(F1' - F1'')^2 + (F2' - F2'')^2}$$

The following example shows how to implement/and demonstrate the E. Distance calculation between /i:/ and /ɪ/ of one participant with normalized values :

$$\begin{aligned}d &= \sqrt{(F1' - F1'')^2 + (F2' - F2'')^2} = \\ &= \sqrt{196 + 1296} = \sqrt{1492} = 38.6\end{aligned}$$

The next sections include data analyses, results, and a final summary

## 6.5 Analysis<sup>5</sup>

### 6.5.1 Vowel distance analyses (F1-F2)

Euclidean distances  $d = \sqrt{(F1' - F1'')^2 + (F2' - F2'')^2}$  were obtained to calculate the degree of separation between the pair of tokens examined. These analyses were conducted for each speaker with the normalised mean value for each pair of vowels: /i:/and /ɪ/, /ɪ/ and /e/ and /u:/and /ʊ/. The larger the Euclidean distance the greater the separation between the contrasting vowels produced by each individual speaker across three time points.

## 6.6 Results

This section presents the spectral and temporal results for the vowels examined. The outcomes include the rates of progress across members of the group in terms of Euclidean distance (ED) separation, F1 and F2 individual changes, followed by the vowel duration results between (/i:/and /ɪ/and /u:/and /ʊ/). High-and low-performing individuals are identified via reference to their ascending or falling values with respect to RP normative values stated in Chapter 5.

---

<sup>5</sup> The formant and duration analyses for the list of words are the same presented in the previous chapter 5.

### 6.6.1 Vowel separation results

The following results present the degree of separation between the vowel contrast /i:/ and /ɪ/, /ɪ/ and /e/, and /u:/and /ʊ/ obtained after calculating the ED (see Appendix F). The results were divided to evaluate the level of changes and the distribution/spread of the vowel pairs at the three times tested. The arbitrary division was based on the ED obtained from the reference data points for RP speakers<sup>6</sup>. The rationale for this was to establish mean reference values to compare and contrast with the results of the present data. For the /i:/ and /ɪ/ vowels, the average for the RP data was 250 Hz; for the /u:/and /ʊ/ vowels, the average was 110 Hz; and for the /ɪ/ and /e/ vowels the reference value was 250 Hz. It is acknowledged that, as these are the average group values, therefore some RP speakers in the study from which they are derived will have ED values for each of these pairs that fall above and below the group means. Unfortunately, the author (Deterding, 1990 as cited in Deterding, 1997) does not provide the range of individual means, and so it is not possible to determine whether participants in the present study produced ED values that fell within the *range* for native L1 speakers. The most that can be ascertained is whether they approached RP group norms. In order to look at differences across time at the individual level, three main groups were established to identify speakers who produced changes in formant values (in terms of separation of vowels) and individuals who did not. The groups were named: (a) *Moderate/static* (b) *Substantial/large movement* and (c) *Backward movement*.

a) *Moderate/static* group: this group was made of participants who did not produce changes in formant values (vowel separation) above the thresholds established for each vowel pairs (e.g., results below 250 Hz for /i:/ and /ɪ/, below 110 Hz for /u:/and /ʊ/, and below 250 for /ɪ/ and /e/).

b) *Substantial/large movement* group: participants assigned to this group were the ones who showed changes in formant values above the thresholds established for each vowel pair or maintained their formant values above the RP norms (e.g., results above 250 Hz for /i:/ and /ɪ/, above 110 Hz for /u:/and /ʊ/, and above 250 Hz for /ɪ/ and /e/).

---

<sup>6</sup> The values were taken from Deterding (1990 as cited in Deterding, 1997) to then calculate the E. distance between the values examined.

c) *Backward movement* group: this group consisted of participants who produced formant values (in terms of separation of vowels) above the thresholds set for the vowel pairs but at some testing point (T2 or T3) produced values below the thresholds.

### 6.6.2 /i:/ and /ɪ/ vowels results

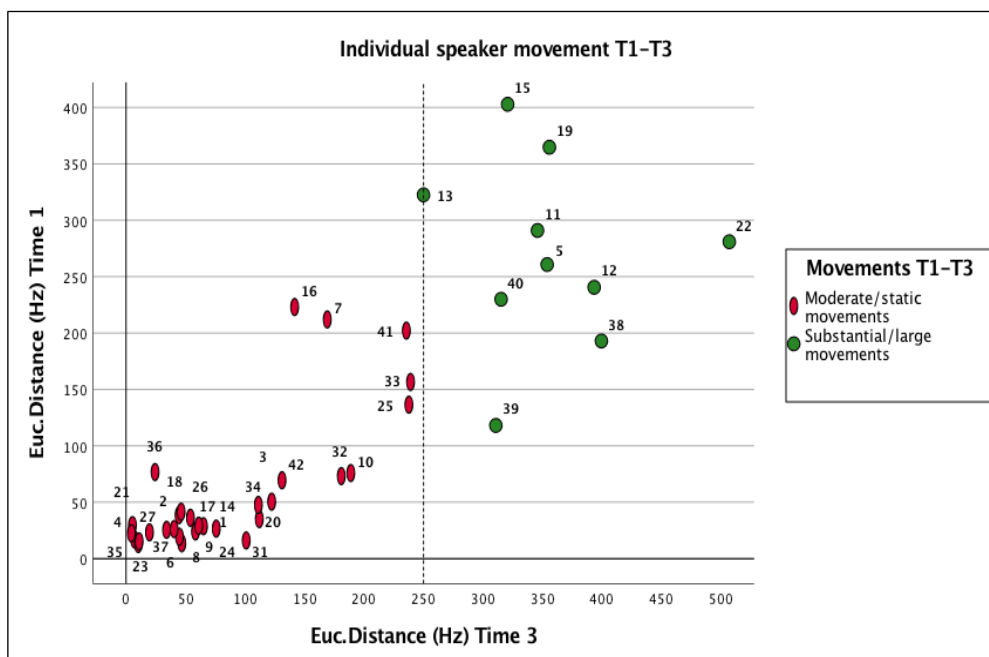
This section shows the results for /i:/ and /ɪ/ separation. Overall, the results show that most of the participants [ $\approx 70\%$ ] maintained their separation boundaries between /i:/ and /ɪ/ below the 250 Hz across times point. The speakers did not sufficiently show distance between /i:/ and /ɪ/.

The results from (T1 to T2) show that 28 low-performing speakers kept their initial separation under the 250 Hz threshold criterion with an average of 89 Hz. Exceptions were seen only among the six high performers who increased their initial distances and moved over 250 Hz with an average of 321 Hz. It should be noted that six out of 12 speakers retained a distance above the 250 Hz from T1 with a mean of 369 Hz.

By T3, the results show a plateau effect. Thirty low-performing speakers did not distance their pair of vowels above 250 Hz, producing an average distance of 87 Hz. That is, the speakers tended to maintain their separation values from T2. Interestingly, 10 high-performing speakers increased their separation values above the RP reference with an average of 355 Hz. Figure 21 and Table 12 provide an overview of the results obtained from T1-T3.

**Figure 21**

*Vowel distance variations between /i:/ and /ɪ/ from Time 1-3*



**Table 12**

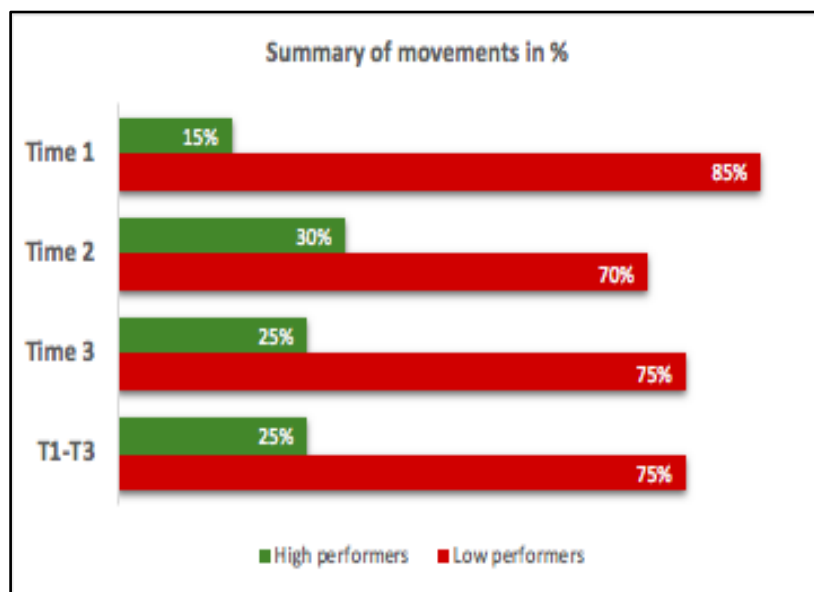
<i>Summary T1-T3</i>	N	%
a) Moderate/static movement (low performers)	30	75%
b) Substantial/large movement (high performers)	10	25%

Overall, in a year (T1-T3), the results indicate that 30 low-performing speakers did not increase their vowel separation and maintained their values with an average of 87 Hz. Minimal variations in the number of speakers per group were observed across times. By contrast, six high-performing speakers kept their initial T1 values above the RP norm, and four speakers developed the contrast at the end of the year.

Figure 22 displays a summary of the results across times as percentages.

## Figure 22

*Summary in percentages about the proportion of speakers who fell into either of the two categories*



In the section that follows, the results are presented in terms of F1 and F2 separately, in order to establish whether the distance of vowels was produced because participants made a distinction between F1 and F2 separately or both formants altogether.

### 6.6.3 F1 and F2 results

The results set out above were obtained based on the EDcalculation, F1 and F2 merged. Considering the outcomes, it was important to look in more detail at what may have caused the distance - or the lack of it - between the pair of vowels under examination, i.e., if there were changes, were these mainly brought about by shifts in F1 or F2? The patterns are shown below<sup>7</sup>.

In terms of F1 by T2, /i:/ results show a divided outcome, half of the values obtained for /i:/ increased resulting in a lower position (376 Hz), while half of the values decreased resulting in a higher position (324 Hz). However, F1 values for /i/ show a tendency (30 speakers) of lowered values, resulting in a higher position than at T1 (426 Hz).

---

<sup>7</sup> I refer to the central position for /i:/ when there is a backward movement in F2 (with consideration to its position as a front vowel). And I refer to the central-back position for /i/ when there is a backward movement in F2 (with consideration to its position as a front vowel).

Similar changes in F2 values were found for /i:/. Almost half of the speakers (22) reported a decreased F2 with the vowel moving towards a less fronted position by T2 (2422 Hz).

Eighteen speakers reported an increased F2 with the vowel moving towards a more fronted position (2482 Hz).

Changes in F2 values found in /ɪ/ show that the majority of the speakers (32) produced a decreased F2, moving to a less fronted position by T2 (2138 Hz).

**Table 13**

*Overall results showing the number of speakers producing higher or lower F1 values from Time 2-3*

<b>Vowel</b>	<b>N° sp</b>	<b>F1</b>	<b>Vowel</b>	<b>N°sp</b>	<b>F1</b>
/i:/	19	Higher values (aveg: 350Hz)	/ɪ/	13	Higher values (aveg: 421 Hz)
	21	Lower values (aveg: 335 Hz)		27	Lower values (aveg: 383Hz)

**Table 14**

*Overall results showing the number of speakers producing higher or lower F2 values from Time 2-3 F2*

<b>Vowel</b>	<b>N° sp</b>	<b>F2</b>	<b>Vowel</b>	<b>N°sp</b>	<b>F2</b>
/i:/	26	Increased values (fronted) (aver: 2599Hz)	/ɪ/	34	Increased values (fronted) (aver: 2406Hz)
	14	Decreased values (central) (aver: 2505Hz)		6	Decreased values (central) (aver: 2209Hz)

As shown in Table 13, for /i:/, F1 values show a steady outcome (from T2 to T3). However, for /ɪ/ a different trend was observed. Twenty-seven speakers produced lower values,



articulating the vowel in a higher position (by T3), and 13 speakers produced it in lower position.

In terms of F2, the number of speakers producing both vowels /i:/ and /ɪ/ with higher values increased, thus, the results showed more fronted vowels compared to T2.

In short, it can be said that at the end of the year, results for /i:/ show that half of the speakers produced a higher fronted vowel, and the other half a lower-fronted vowel; however, the /ɪ/ was produced by most of the speakers as a lower and fronted vowel.

From the above results, it is important to highlight that in terms of F1 by T3, 15 speakers who produced a higher /i:/ also produced a higher /ɪ/. And 21 speakers who produced a fronted /i:/ also did it for /ɪ/.

#### **6.6.4 /u:/ and /ʊ/ results**

The following results present the degree of separation between /u:/ and /ʊ/ vowels based on the ED of the normalised data. Results are presented in groups; the division of the groups was established at 110 Hz, based on the reference data points from RP previously presented.

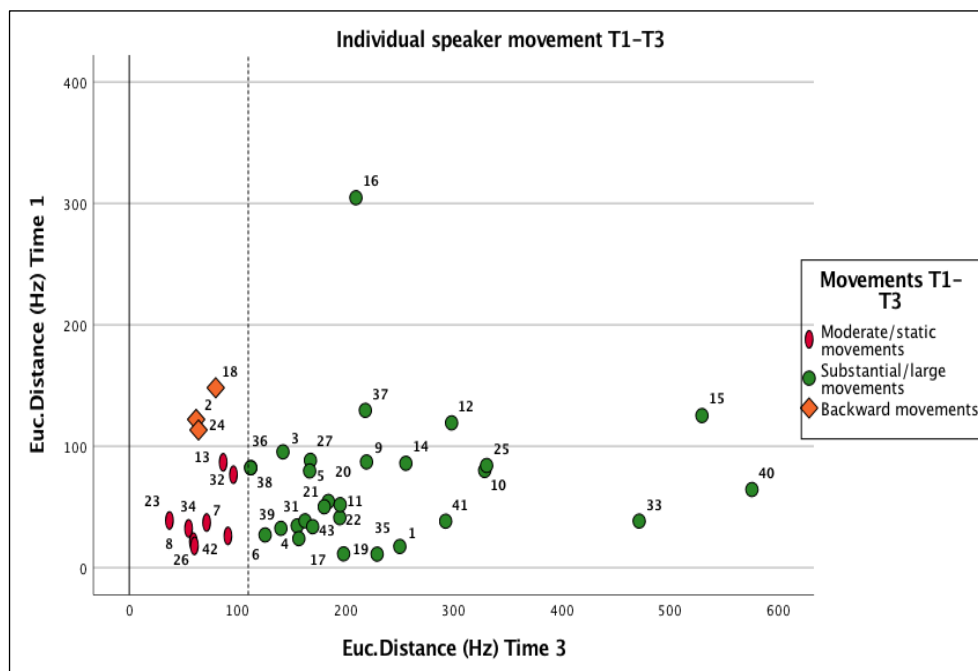
Results at T1 show that seven high- performing speakers produced a distinction between the pair above the RP band (151 Hz). Interestingly, by T2, four speakers who initially showed a distinction between the vowels (above 110 Hz) reversed their values under the threshold (46 Hz), producing closer vowels - below the British English norm. By T2, 25 low- performing speakers maintained their preliminary degree of vowel separation below 110 Hz with an average of 48 Hz. By T2 eight speakers increased their separation vowel difference above the English level (152 Hz), and three speakers maintained their distinction between values from T1 to T2.

By T3, the number of speakers who produced values above 110 Hz increased. Twenty-one speakers increased their separation values, with a mean of 239 Hz, and eight speakers kept their T2 values above the RP level (217 Hz). Finally, a total of 29 high- performing speakers produced a distinction between the pair greater than 110 Hz. Exceptions to this trend were shown by eight speakers who did not produce a distinction between the vowels by T2 and T3 (68 Hz) above the threshold, and two participants who reversed their outcomes (from T2) to less distinct realisations by T3 (75 Hz).

Figure 23 and Table 15 provide an overview of the results obtained at T1 and T3.

**Figure 23**

*Vowel distance variations between /u:/ and /o/ from Time 1-3*



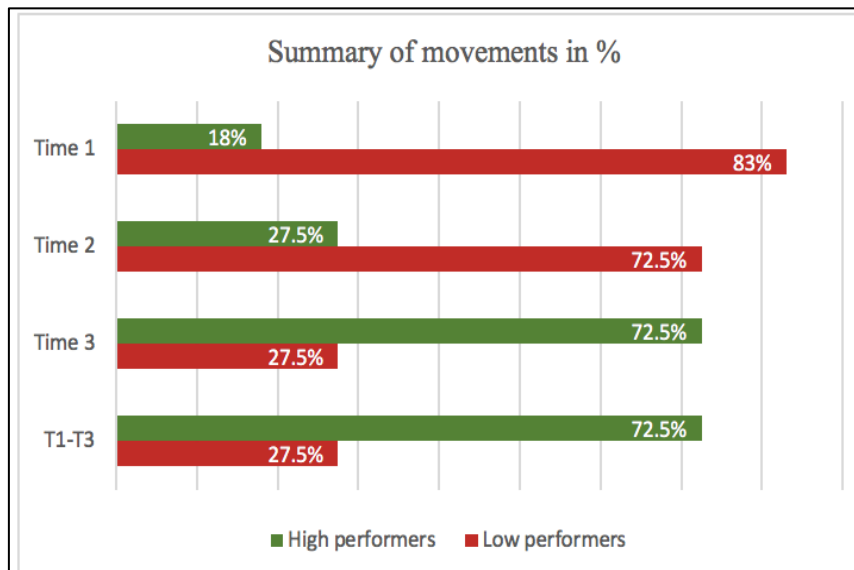
**Table 15**

<i>Summary T1-T3</i>	<b>N</b>	<b>%</b>
a) Moderate/static movement (low performers)	8	20%
b) Substantial/large movement (high performers)	29	72.5%
c) Backward movement (low performers)	3	7.5%

Overall, in a year (T1-T3), the results indicate that the number of high performing speakers increased by T3. At the end of the year, 25 high performing participants increased their separation vowel difference above the RP set band of 110 Hz, their values changing from 53 Hz to 220 Hz. Eleven low performing speakers did not increase the distance between their pair of vowels and kept their values under the threshold (69 Hz), during the course of the year. Figure 24 shows a summary of the results in percentages.

**Figure 24**

*Summary in percentages about the proportion of speakers who fell into either of the two categories.*



In order to determine whether the distance of vowels already presented was obtained because participants were able to make a distinction in F1 and F2 separately or because they changed both formants together, the next section shows the results in terms of individual formants (F1, F2).

### **6.6.5 F1-F2 results**

Results for the acoustic changes in F1 and/or F2 for /u:/ and /ʊ/ are set out below. Generally, the F1 values for both vowels, especially by T2, tended to have a similar number of speakers producing higher and lower vowels. In terms of the F2 values, both values had a tendency to increase over the year.

Results by T2 show that almost half of the participants presented F1 values for /u:/ and /ʊ/ higher than at T1 with an average for /u:/ of 422 Hz and /ʊ/ of 463 Hz. The other half of the speakers presented the F1 values lower than at T1 with an average for /u:/ of 372 Hz and /ʊ/ of 415 Hz.

In terms of F2, both vowels were produced - by the majority of the speakers - with lower values than at T1 with an average for /u:/ of 976 Hz and /ʊ/ of 1019 Hz.

**Table 16**

*Overall results showing the number of speakers producing higher or lower F1 values from Time 2-3*

<b>Vowel</b>	<b>N° sp</b>	<b>F1</b>	<b>Vowel</b>	<b>N°sp</b>	<b>F1</b>
/u:/	17	Higher values (avg: 412Hz)	/ʊ/	27	Higher values (avg: 538Hz)
	23	Lower values (avg: 365)		13	Lower values (avg: 406Hz)

**Table 17**

*Overall results showing the number of speakers producing higher or lower F2 values from Time 2-3*

<b>Vowel</b>	<b>N° sp</b>	<b>F2</b>	<b>Vowel</b>	<b>N°sp</b>	<b>F2</b>
/u:/	39	Increased values (fronted) (aver: 1519Hz)	/ʊ/	40	Increased values (fronted) (aver: 1331Hz)
	1	Decreased values (retracted) (aver: 843hz)		0	Decreased values (retracted)

Clear outcomes are seen in Tables 16 - 17. F1 for /u:/ presents a similar division in the number of the speakers who produced higher and lower values at T2. However, F1 for /ʊ/ shows an increase in the number of speakers producing higher values by T3.

In terms of F2, striking results were found for both /u:/ and /ʊ / vowels. All of the speakers (with only one exception for /u:/) increased their F2 values, shifting their initial retracted vowel towards a more fronted variant by T3.

From the results described above, it is clear that by T3 for /u:/, the same 22 speakers who produced lower F1 values also produced higher F2 values. For /ʊ/ by T3, the 27 same

speakers who produced the vowel with higher F1 also did the same with higher F2 values, and the same 13 speakers who produced lower F1 values also produced higher F2 values. Finally, in terms of F2 by T3, all of the speakers produced a more fronted /u:/ and /ʊ/ than at T1.

To sum up, it is possible to say that regarding F1, by T2 the vowel /u:/ generally tended to move towards /ʊ/ (higher F1) but, by T3, a bigger separation between the vowels was observed, whereby /u:/ exhibited lower F1 values (365 Hz) and /ʊ/ followed the opposite tendency (538 Hz). In terms of F2 by T3, both vowels were produced in a more fronted position. /u:/ obtained a mean of 1519 Hz, while /ʊ/ obtained a mean of 1331 Hz.

I will refer back to these values and the comparison with the RP data in the summary section.

#### **6.6.6 /ɪ/ and /e/ results**

The results for the separation between /ɪ/ and /e/ are presented below. The division of the groups was settled at 250 Hz, the threshold being established based on the reference points from the RP data presented in Chapter 5.

Overall, participants exhibit few variations in the separation of the vowels across Time. Twenty-four speakers produced a distinction between the pair at T1 (312 Hz) above the threshold. Twenty-three speakers made it by T2 (332 Hz), and 23 speakers kept a distinction between the vowels by T3 (340 Hz).

The results by T2 show that out of the 24 speakers who initially produced a distinction between the pair of vowels above the threshold, 20 maintained it by T2 with an average of 332 Hz. Three speakers improved their vowel separation above the RP values with an average of 290 Hz. Four speakers who had a distinction at T1 above the RP norm had not developed one by T2 (174 Hz), and 12 low performing speakers who did not distance their pair of vowels at T1 did not develop one by T2.

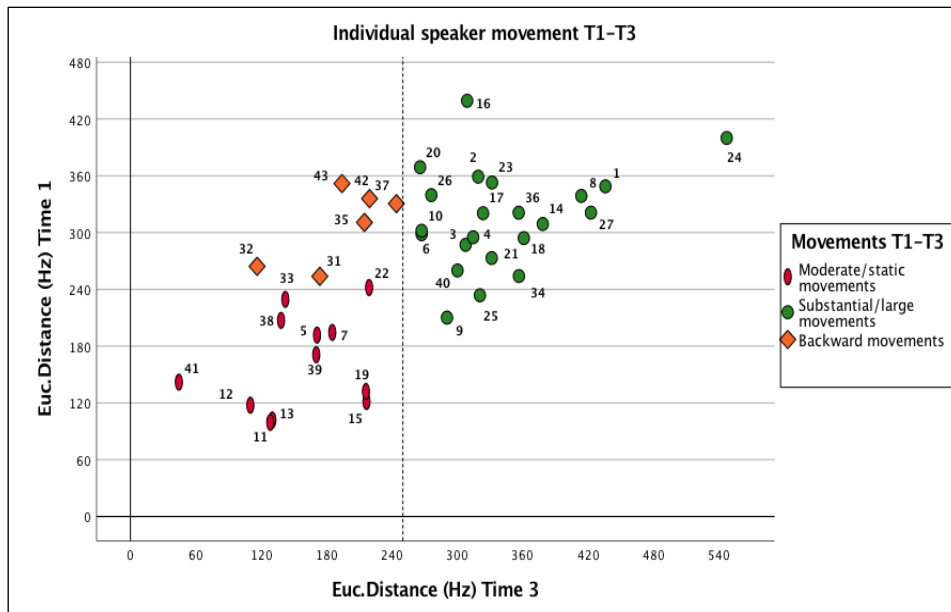
The results from T2 to T3 are steady; there were no major changes occurring from T2 to T3. By T3, a total of 23 high performing speakers had a distinction between this pair (340Hz). Four participants who did not have a distinction above the threshold by T2 developed one by

T3 (327 Hz) and a total of 17 low performers did not make progress to distance their vowel pair (165 Hz).

Figure 25 and Table 18 show an overview of the results obtained from T1 and T3.

**Figure 25**

*Vowel distance variations /ɪ/ and /e/ from Time 1-3*



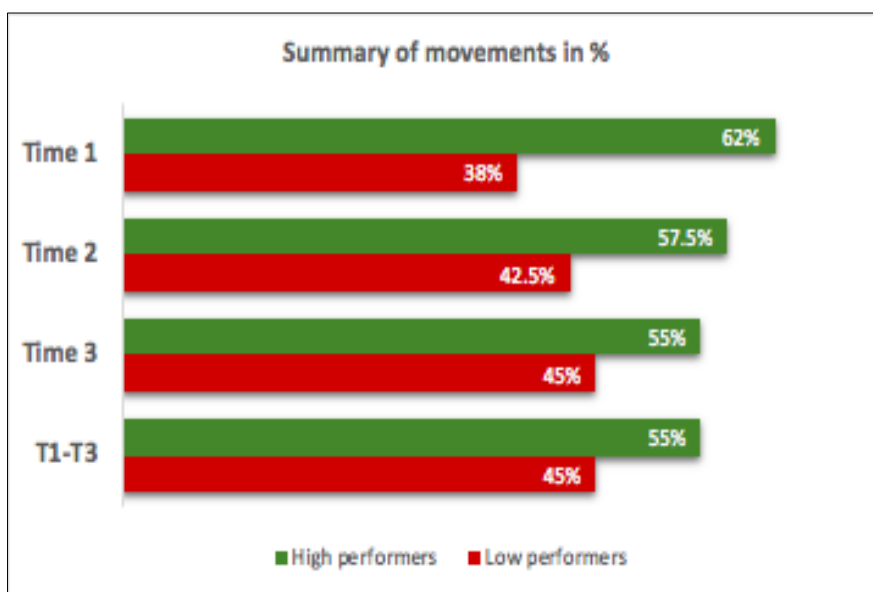
**Table 18**

<i>Summary T1-T3</i>	N	%
a) Moderate/static movement (low performers)	12	30%
b) Substantial/large movement (high performers)	22	55%
c) Backward movement (low performers)	6	15%

Overall, the previous results show a gradual increase in the distance between /ɪ/ and /e/ between T1 (312 Hz) and T3 (340Hz). Only two speakers who did not distance their pair of vowels above the threshold at the start of the year had progressed to making one by the end (305 Hz).

## Figure 26

*Summary in percentages about the proportion of speakers who fell into either of the two categories*



In the section that follows results are presented in terms of F1 and F2 separately.

### 6.6.7 F1-F2 results

Generally, over a year, it can be said that F1 values for /ɪ/ increased from 301 Hz to 334 Hz and /e/ decreased from 645 Hz to 554 Hz. Concerning F2, /ɪ/ showed a tendency of an overall decrease in the frequency of the second formant (from 2608 Hz to 2571 Hz), and /e/ showed a tendency of an overall increase in the frequency of F2 (1994 Hz to 2266 Hz), indicating a production with a more fronted tongue position than at T1<sup>8</sup>.

By T2, F1 results for /ɪ/ show that 30 speakers produced higher values compared to T1 going from 301 Hz to 426 Hz. For /e/, 20 speakers produced lower values, going from 645 Hz to 588 Hz, and 20 speakers produced higher values with an average of 676 Hz.

In terms of F2 almost the same number of speakers produced lower values for both vowels.

---

<sup>8</sup> I refer to the central-back position for /ɪ/ when there are low values in F2. The same applies for /e/.

**Table 19**

*Overall results showing the number of speakers producing higher or lower F1 values from Time 2-3*

<b>Vowel</b>	<b>N° sp</b>	<b>F1</b>	<b>Vowel</b>	<b>N°sp</b>	<b>F1</b>
/i/	13	Higher values (avg: 421 Hz)	/e/	16	Higher values (avg: 669Hz)
	27	Lower values (avg: 383Hz)		24	Lower values (avg: 587Hz)

**Table 20**

*Overall results showing the number of speakers producing higher or lower F2 values from Time 2-3*

<b>Vowel</b>	<b>N° sp</b>	<b>F2</b>	<b>Vowel</b>	<b>N°sp</b>	<b>F2</b>
/i/	34	Increased values (fronted) (aver: 2406Hz)	/e/	30	Increased values (fronted) (aver: 2011Hz)
	6	Decreased values (central- back) (aver: 2209Hz)		10	Decreased values (retracted) (aver: 1901Hz)

By T3, in terms of F1 and F2, the above tables show a clear pattern for both vowels.

F1 results for /i/ and /e/ show that most of the speakers shifted their vowel production (T2-T3) from higher values to lower ones. For /i/, 27 speakers decreased their values from 426 Hz to 387 Hz, and for /e/, 24 speakers decreased their F1 values from 676 Hz to 587 Hz.

In terms of F2, by T3, most of the speakers changed their production of both vowels towards a more fronted position compared to the previous time. The production of /i/ increased from 2138 Hz to 2406 Hz, and /e/ increased from 1836 Hz to 2011 Hz.

It is important to note that by T2, the same 16 speakers who produced higher F1 values for /i/ also did so for /e/. The 23 speakers who produced lower F2 values for /i/ also produced lower F2 values for /e/. Additionally, by T3, 16 out of 40 speakers who produced lower F1 values



for /ɪ/ also did so for /e/, and the 26 speakers who increased their F2 values for /ɪ/ also increased their values for /e/.

In summary, /ɪ/ tended to move to a lower and fronted position and /e/ was inclined towards a higher and fronted position, possibly owing to an adjustment towards native speaker norms (a vowel close to cardinal 3).

Together these results provide important insights into the progress that participants have made in developing the contrast of English vowel pairs /i:/ and /ɪ/, /ɪ/ and /e/ and /u:/ and /ʊ/ with respect to spectral features. Also, it has been possible to observe if the progress or changes that speakers made regarding vowel contrasts have been done in terms of F1 and F2 separately or both formants together.

Despite the degree of separation between the vowel pairs examined participants generally showed some formant differentiation. That is, the overall realisation of vowels by the Spanish speakers tends to follow the British English tendency namely: /u:/ produced in a higher position than /ʊ/, /i:/ produced in a higher position than /ɪ/, and /e/ produced lower than /ɪ/. These results and their comparison to the reference values from the contemporary RP data will also be referred in the final summary section below.

### **6.7 /u:/ and /ʊ/ fronting results**

As discussed in section 6.4, the fronting of the two high- back vowels /u:/ and /ʊ/, especially in Southern British English, has been observed in recent years. To determine if this trend was produced by the participants of this research while reading citation forms, the fronting of /u:/ and /ʊ/ was analysed in general terms. The rationale for considering if speakers fronted their vowels was based on the results obtained at T1 (movements towards fronted or retracted position were observed starting from there).

Only F2 normalised values (18 in total) were considered when reporting the possible changes of the vowels across time. Based on the results, three groups were created to show the formant variations: Fronting, Backing and the Same (no changes in F2 between times) groups. A threshold of  $\pm 20\text{Hz}$  was established to consider changes among groups. The reasoning for the

threshold was based on Ladefoged (1996, p.21) who indicates that human ears cannot detect (low) frequencies below 20 Hz. (see Appendix G for Figures).

*Fronting* group: participants assigned to this group had a fronting F2 movement of +20 Hz by Time to Time.

*Backing* group: participants assigned to this group retracted their vowels +20 Hz from Time to Time.

*Same* group: participants assigned to this group had minimal F2 variation between Times (1-20 Hz).

The results of /u:/ and /ʊ/ fronting are presented in the following sections.

### **6.7.1 /u:/ - fronting**

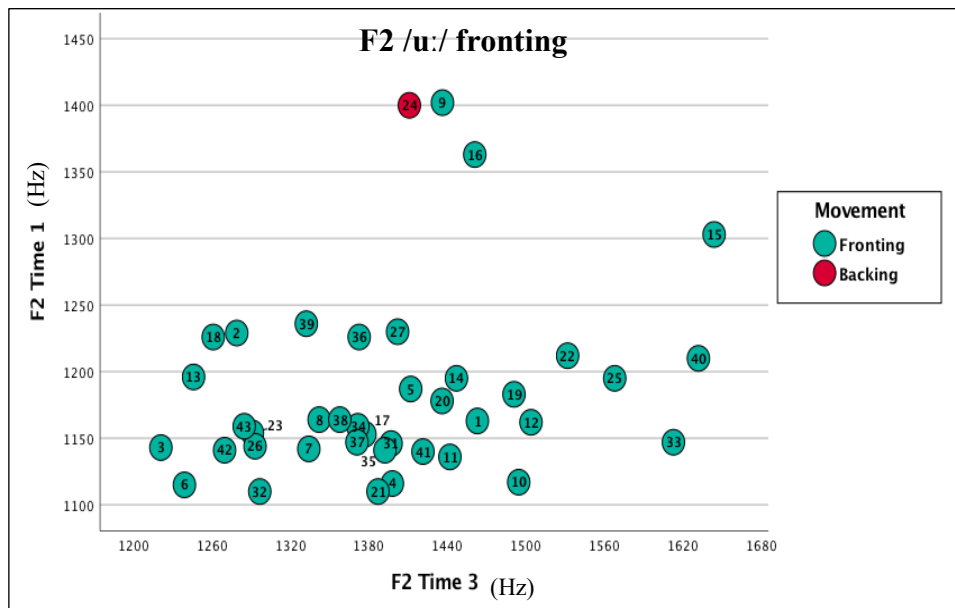
The results show that overall realisation at T1 was 1186 Hz. The results by T2 show that 13 speakers (who initially produced a more fronted /u:/ vowel, avg.:1237 Hz) retracted their F2 vowel (avg:1165 Hz), and 13 speakers increased their original F2 values from an average of avg 1152 Hz to an average of 1214. Fourteen speakers maintained their F2 values (avg:1169 Hz). Interestingly, by T3 all of the speakers fronted their production of /u:/ from an average of 1182 Hz to 1398 Hz. In other words, the majority of the speakers did show /u:/ fronting.

An observable gradual increment in the formant values was produced during the year by all the speakers with the exception of one (Figure 27). Overall, the group realisation started at 1186 Hz. By T2, for those who increased their values it was 1214 Hz, and by T3, it was 1398 Hz.

Figure 27 shows the overall performance from T1-T3.

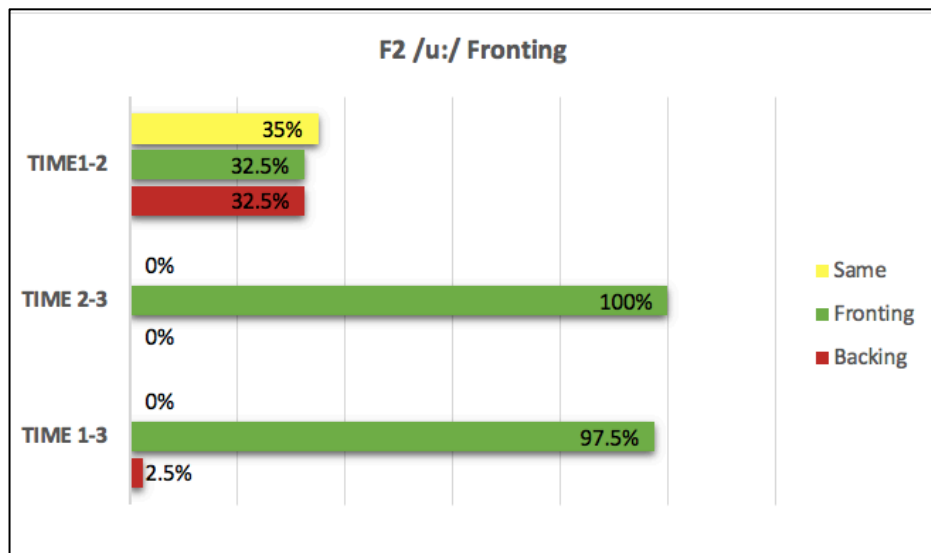
**Figure 27**

*Changes in the F2 values for /u:/ between T1 and T3*



**Figure 28**

*Summary in percentages about the proportion of speakers who fell into either of the three categories*



### 6.7.2 /ʊ/ - fronting

The results for the production of /ʊ/ show that at T1 the vowel was realised with an average of 1174 Hz. This retracted production changed by T2. Twenty-seven speakers who began the year as non-fronters increased their F2 values from 1151 Hz to 1219 Hz. Five participants decreased their initial F2 values from 1285 Hz to 1153 Hz, producing a more retracted vowel, and eight speakers did not produce noticeable changes between these two times and maintained the values with an average of approximately 1188 Hz.

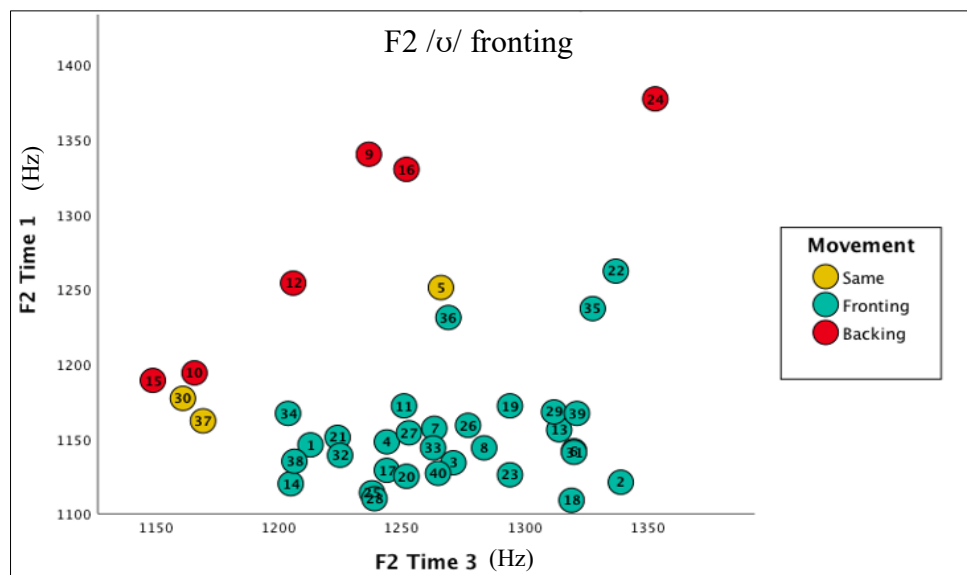
By T3, 25 speakers increased their F2 values for /ʊ/ from 1187 Hz to 1282 Hz, producing a more fronted vowel. Eleven speakers kept their values from T2 (1226 Hz), and four speakers retracted their second formant from 1262 Hz to 1198 Hz.

These results show that even though there was a backward and forward movement among the speakers, the majority increased their F2 values and produced a more fronted vowel than at T1.

Figures 29 & 30 show a summary of the overall results.

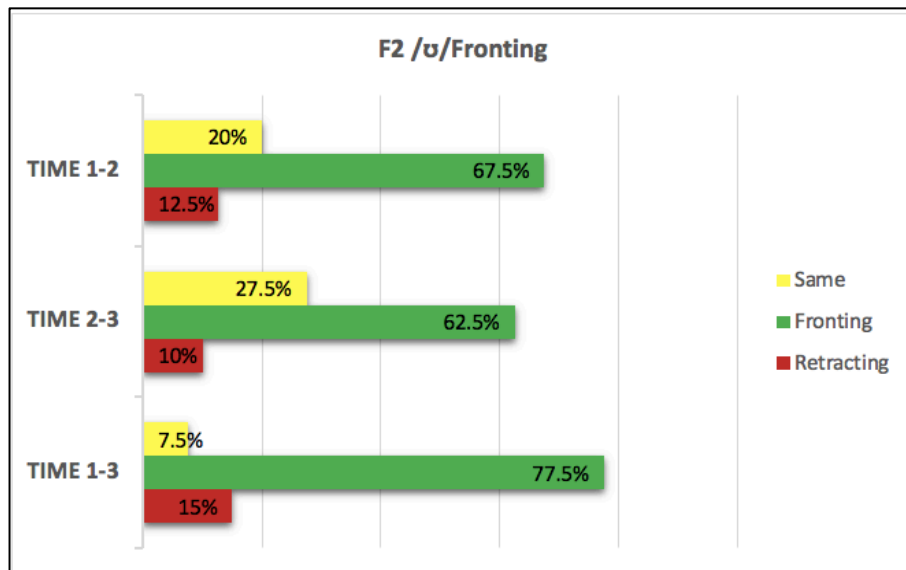
**Figure 29**

*Changes in the F2 values for /ʊ/ between T1 and T3*



**Figure 30**

*Summary in percentages about the proportion of speakers who fell into either of the three categories*



A summary of the formant findings is presented below, after the duration analyses in the chapter summary.

### **6.8 Vowel duration results**

As discussed in Chapter 2 the length distinction for long vowels, in this case /i:/ and /u:/, should only be considered under certain conditions because native speakers' production of /i:/ and /u:/ is not always longer than its contrast /ɪ/ and /ʊ/ in all phonological contexts.

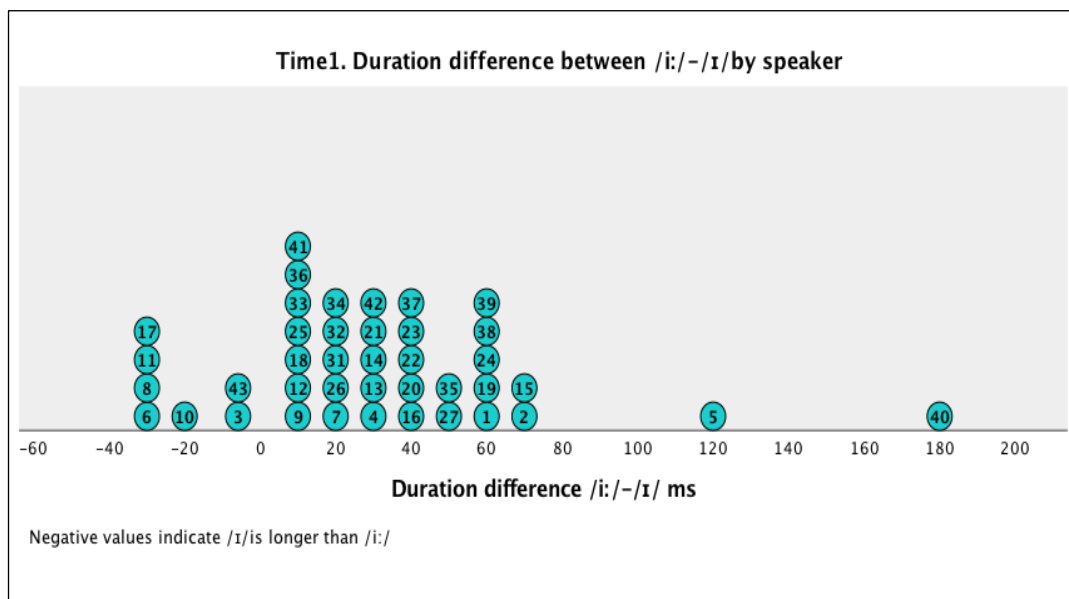
Therefore, progress towards native likeness was measured by looking at vowels followed by a voiced consonant only. However, by T1 and solely for /i:/ (as in Chapter 5), it also examined the length data for a condition where the vowel was followed by a voiceless consonant, because if a durational distinction occurred under this circumstance, it might suggest that participants over-extended the rule of producing a longer vowel if the syllable is closed by the voiced consonant.

### 6.8.1 /i:/ and /ɪ/ vowels

The quantity distinction between the /i:/ and /ɪ/ vowels are presented in this section. Based on the outcomes, two groups were formed: a) positive values (correct length) and b) negative values<sup>9</sup> (incorrect length).

**Figure 31**

*Duration difference between /i:/ and /ɪ/ at Time 1 (voiceless final consonants)*



**Table 21**

*Summary T1*

<i>Summary T1</i>	<b>N</b>	<b>%</b>
a) Positive separators	33	82.5%
b) Negative separators (/ɪ/ longer than /i:/)	7	17.5%

Figure 31 shows that, at T1, 33 speakers produced the vowel /i:/ longer than the /ɪ/ despite the phonological context in which the /i:/ was presented. The average length for the production of /i:/ was 427 ms and for /ɪ/ was 372 ms. Seven participants produced the correct length distinction (just by means of /i:/ being 22 ms longer than /ɪ/) with an average for /i:/ of 454 ms

<sup>9</sup> Negative values indicate that /ɪ/ was produced longer than /i:/ when it should have been shorter.

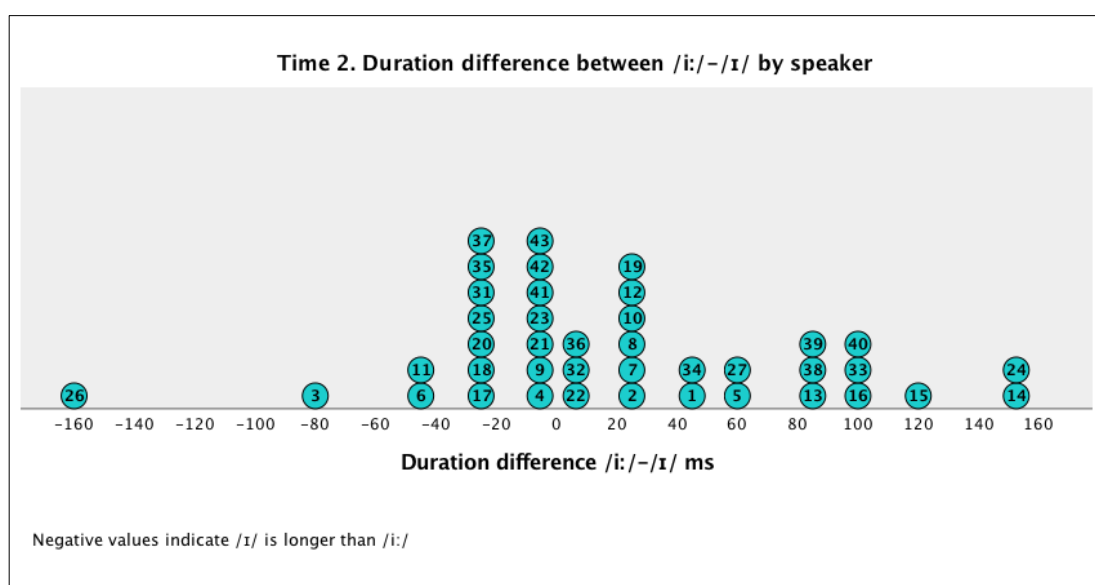
and /ɪ/ of 476 ms. Although the average direction of the duration of these two vowels is 'somehow correct' it does not approach the British English norms.

Thus, it can be said that at T1, most of the participants, with the exception of seven, over-extended the consonant length rule, and produced a longer /i:/, even though the vowel was followed by a voiceless consonant.

Figures 32-33 present the results of vowels for T2 and T3 which were followed by a voiced consonant.

**Figure 32**

*Duration difference between /i:/ and /ɪ/ at Time 2 (Voiced final consonants)*



**Table 22**

*Summary T2*

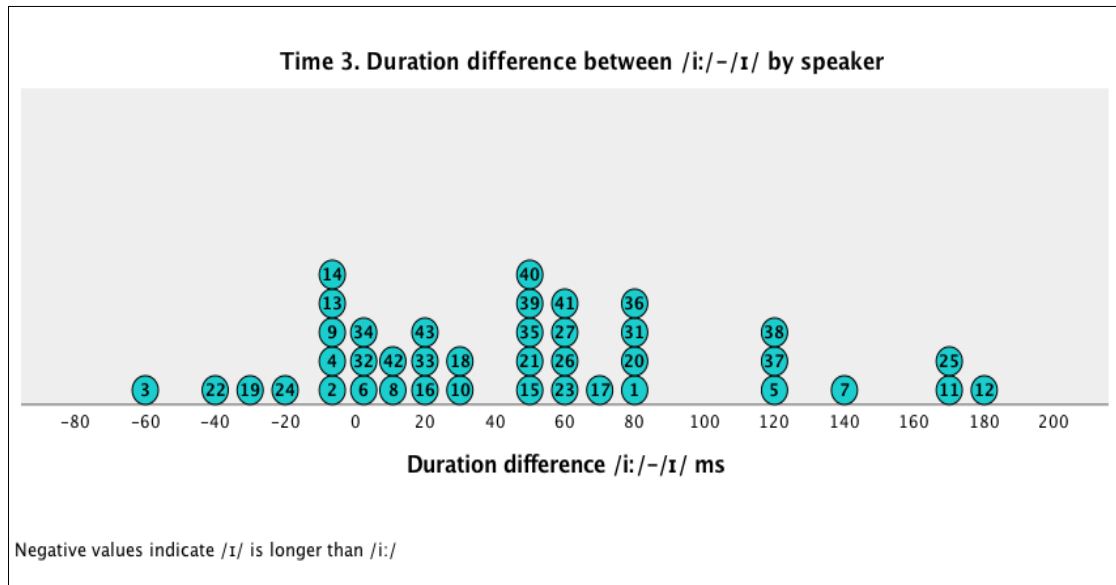
<i>Summary T2</i>	N	%
a) Positive separators	22	55%
b) Negative separators (/ɪ/ longer than /i:/)	18	45%

Interestingly, by T2 (Figure 32) the results revealed a rather uniform distribution of speakers producing the correct and incorrect length distinction between the pair of vowels. Twenty-two speakers produced the /i:/ longer than the /ɪ/ with the average values for /i:/ of 484ms and /ɪ/

442ms. On the contrary, 18 speakers who produced the reversed length distinction did it with an average for /i:/ of 380 ms and /ɪ/ with an average of 491 ms.

**Figure 33**

*Duration difference between /i:/ and /ɪ/ at Time 3*



**Table 23**

*Summary T3*

<i>Summary T3</i>	<b>N</b>	<b>%</b>
a) Positive separators	31	77.5%
b) Negative separators (/ɪ/ longer than /i:/)	9	22.5%

As shown in Figure 33, by T3, 31 speakers produced positive values with a mean for /i:/ of 451 ms and /ɪ/ 383 ms, the exception being produced by nine speakers who maintained their production of /ɪ/ as longer than /i:/ (492 ms and 471 ms respectively). Fifteen speakers who by T2 produced a longer /ɪ/ shifted their tendency by T3, and six speakers who by T2 managed the correct length distinction, by T3 showed reverse outcomes.

Overall, the results show an increment in the percentage of speakers who produced ‘the correct’ vowel length distinction from T2 (55%) to T3 (78%). However, these results (/i:/ 451 ms and /ɪ/ 383 ms), although they are in the direction associated with native English, do not



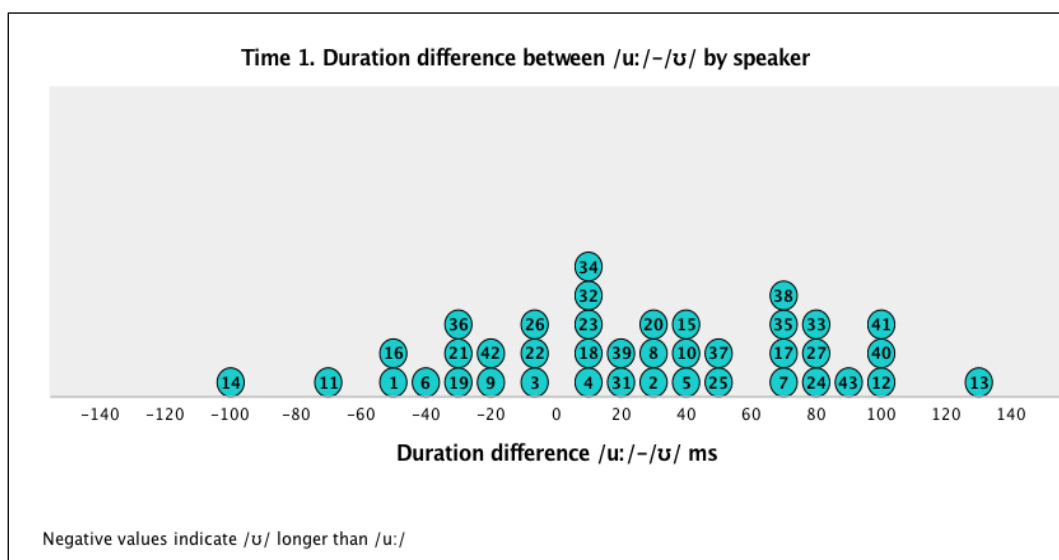
show length distinction between the pairs since the difference involved only 68ms. And it does not indicate that the speakers shift towards the British English norm, which established /i:/ with a length of 293 ms and /ɪ/ with 139 ms.

### 6.8.2 /u:/ and /ʊ/ vowels

This section shows the results for the length distinction (ms) between /u:/ and /ʊ/. Only vowels followed by voiced consonants were analysed.

**Figure 34**

*Duration difference between /u:/ and /ʊ/ at Time 1*



**Table 24**

*Summary T1*

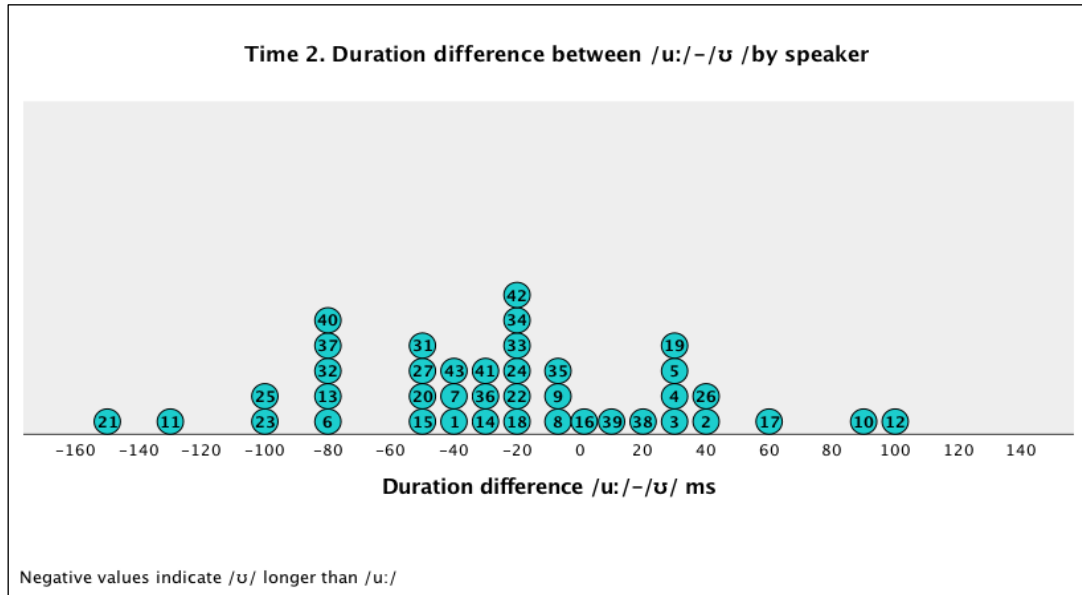
<i>Summary T1</i>	N	%
a) Positive separators	27	67.5%
b) Negative separators (/ʊ/ longer than /u:/)	13	32.5%

Figure 34 shows a clear illustration of the speakers' distribution at the moment of producing the length contrast between /u:/ and /ʊ/. Results show that 27 speakers produced /u:/ longer than /ʊ/ with an average duration for /u:/ of 411 ms and for /ʊ/ of 370 ms. However, 13

participants did not follow this tendency and produced the opposite results, producing /u:/ with 426 ms and /ʊ/ with 432 ms.

**Figure 35**

*Duration difference between /u:/ and /ʊ/ at Time 2*



**Table 25**

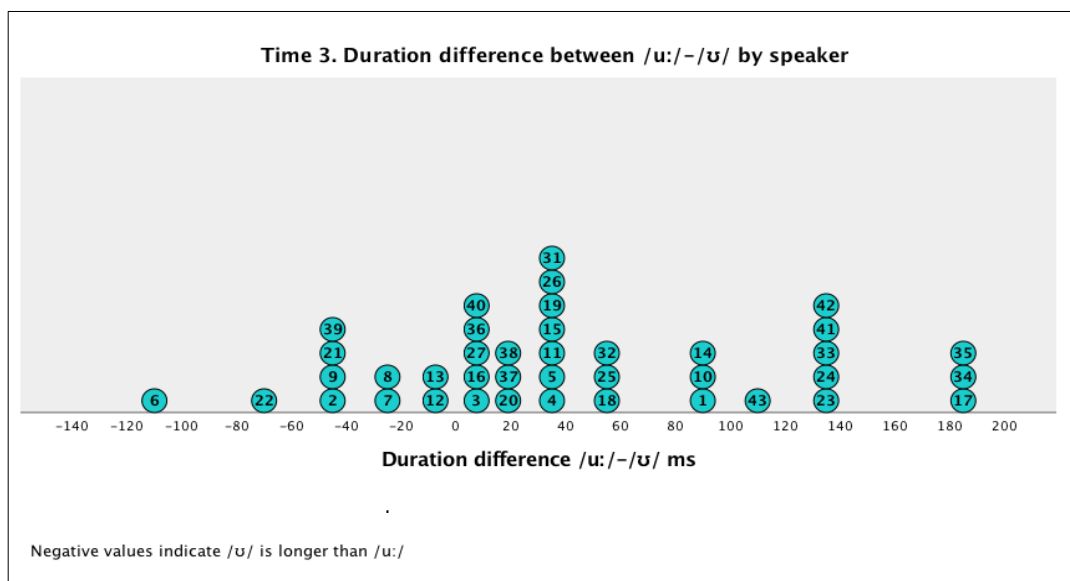
*Summary T2*

<i>Summary T2</i>	N	%
a) Positive separators	12	30%
b) Negative separators (/ʊ/ longer than /u:/)	28	70%

Contrary to T1, by T2 28 speakers, as seen in Table 25, produced the reversed vowel values. The results show an increase in the length of /ʊ/ with an average of 431 ms, and a decrease in the length of /u:/ with an average of 384 ms. Only 12 speakers managed to produce positive values with an average for /u:/ of 512 ms and for /ʊ/ of 471 ms.

**Figure 36**

*Duration difference between /u:/ and /ʊ/ at Time 3*



**Table 26**

*Summary T3*

<i>Summary T3</i>	N	%
a) Positive separators	30	75%
b) Negative separators (/ʊ/ longer than /u:/)	10	25%

Results by T3 (Figure 36) show that 21 speakers who previously produced the incorrect length distinction reverted their results to their earlier realisations: that is, they went back to producing /u:/ longer than /ʊ/. By T3, 30 out of 40 speakers produced positive vowel values, /u:/ was produced with a length of 491 ms and /ʊ/ with a length of 482 ms. On the contrary, 10 speakers produced /u:/ with a length of 473 ms and /ʊ/ with a length of 510 ms. Overall, over a year, (T1-T3) 21 speakers improved their vowel duration contrast. However, 10 speakers were not able to produce positive values between the pair. Nevertheless, the fact that 30 speakers managed to produce /u:/ 9 ms longer than /ʊ/ (491 ms vs 482 ms) does not indicate that the speakers shift towards the British English norm, which established /u:/ with a length of 294 ms and /ʊ/ with 142 ms.

To sum up, by comparing the results between /i:/ and /ɪ/ and /u:/ and /ʊ/, it can be seen that by the end of the year some participants were able to produce positive values between both

vowel pairs; however, that only resulted in a produced length distinction in the direction associated with native English. The durational differences in all cases were between 10- 70 ms which does not show participants shifting towards the English norm.

This section has shown the results for the quantity distinction between vowels. The subsequent section moves on to consider the results by providing the final chapter summary.

## **6.9 Summary**

The aim of the present chapter was to determine the degree of separation that each individual speaker achieved across time points between contrasts in the English vowel pairs /i:/ and /ɪ/, /u:/ and /ʊ/ and /e/ with respect to both spectral and temporal features using citation-style speech. The results were compared to the RP reference values specified in Chapter 5 to observe if the participants were shifting or adapting to the British English norms.

### **6.9.1 Formant analysis (vowel quality)**

To summarise individual performance at the beginning and end of the year as measured by F1 - F2 Euclidean distances:

*/i:/ vs /ɪ/*

- Six speakers began the year with a distinction between this pair very close to or above the RP average (over 250 Hz). They still had such a distinction at the end of the year.
- Four speakers who did not have a distinction of this order at the start of the year had progressed to making one by the end.
- The other 30 speakers, who did not have a distinction of this magnitude at the beginning of the year, had not developed one by the end.

### **/ɪ/ vs /e/**

- Twenty-six speakers had a near-native distinction between this pair at the start of the year (over 250 Hz E.D).
- By the end of the year 20 had maintained the distinction.
- Six speakers who had the distinction at T1 had not maintained it by the end of the year.
- Two speakers, who did not have a near-native distinction at the start of the year, had progressed to making one by the end.
- Twelve speakers who did not have a near-native distinction at the beginning of the year had not developed one by the end.

### **/u:/ vs /ʊ/**

- Seven speakers had a near native distinction between this pair at the start of the year (over 110 Hz).
- By the end of the year five had maintained the distinction.
- Two speakers who had the distinction at T1 had not maintained it by the end of the year.
- Also, by the end of the year, 25 speakers who did not have an RP-like distinction at the start of the year, had progressed to making one by the end.
- Eight speakers who did not have such a distinction at the beginning of the year, had not developed one by the end.

### **/u:/-fronting**

To summarise individual performance between the beginning and end of the year as measured by F2 values:

- At T1, the group realisation of /u:/ was produced as a close back rounded vowel in the vicinity of Cardinal Vowel 8 (average F2 for the group = 1186 Hz).
- By the end of the year 39 speakers produced more fronted variants (average F2 for this group = 1398 Hz, i.e., F2 had been increased by 20 Hz or more).
- Only 1 speaker did not have fronting of his/her F2 realisation by the end of the year.

## **/ʊ/-fronting**

To summarise individual performance at the beginning and end of the year as measured by F2 values:

- At T1, the group realisation of /ʊ/ was a retracted one (average F2 for the group = 1174 Hz).
- Thirty-one speakers who began the year as non-fronters were producing more fronted realisations by the end (average F2 at T3 for this group = 1270 Hz, i.e., F2 had been increased by 20 Hz or more).
- Six speakers were producing more retracted realisations by the end of the year (average F2 at T3 for this group = 1227 Hz, i.e., F2 had been reduced by 20 Hz or more).
- The remaining 3 speakers had not changed their initial realisation by the end of the year (average F2 at T3 for this group = 1198 Hz).

In respect of the beginning to end of year patterns for the vowel quality data, it is of interest that, as seen in the preceding sections, for some vowels and pairs, the development is not unidirectional over the year's course. For example, the T2 data show that some speakers had not progressed between T1 and T2, and some had actually regressed, before progressing again towards the native English values by T3.

### **6.9.2 Vowel duration**

- At the beginning of the year, the participants did not make a clear quantity distinction between /i:/ and /ɪ/, and /u:/ and /ʊ/. Some speakers, in fact, produced the short members of these pairs fractionally longer than the long members; others made the distinction in the direction associated with native English, but the durational differences in all cases only involved only 9-70 ms, which means that participants overall did not produce a length difference.
- By the end of the year, the situation remained much the same.

### 6.9.3 General

This chapter has provided evidence of the different degrees of separation that each individual speaker exhibited across time points between contrasts in the English vowel pairs /i:/ and /ɪ/, /ɪ/ and /e/, and /u:/ and /ʊ/ with respect to both spectral and temporal features using citation-style speech.

A discussion of the findings set out in this chapter is deferred until the final Discussion chapter.

Chapter 7 examines whether the quality and durational distinctions derived from the analyses and results presented in this chapter have consequences for how British English native speakers hear Spanish speakers' word productions.

## Chapter 7. Perception by Native British English Speakers

### 7.1 Introduction

The analysis of the Spanish speakers' productions is based on the speakers making distinctions between the vowels /i:/ and /ɪ/, /ɪ/ and /e/ and /u:/ and /ʊ/ on the basis of vowel quality - as measured by formant configurations - and durations. The perception test with native English raters described here is a tool designed to determine whether the quality and durational distinctions derived from analysis of the production data had consequences for how the productions are heard.

Researchers have used different methods and measurements, such as intelligibility, comprehensibility and accentedness, to assess L2 speech production. Munro & Derwing (1995) suggest that these dimensions, although connected, are also autonomous. They distinguish these three speech evaluation criteria in the following way: intelligibility is the degree of understanding of a production; comprehensibility is the level of perceived effort made by the listener to understand a production; and accentedness is the degree of foreign accent in a production that a native speaker is able to perceive. In terms of methodology, these measurements show some differences in the assessment of the L2 production.

Comprehensibility and accentedness dimensions are usually evaluated by using different types of Likert scales, while intelligibility is commonly judged by counting/processing the accurate written responses given by the raters after listening to non-native speech (Munro & Derwing, 1995). However, an alternative viewpoint by Levis (2006) considers intelligibility as broadly close to comprehensibility, claiming that both notions describe how a person understands an utterance or performance that is not exempt from socially biased judgements. Further, Fraser (2003) points out that the listener's expectations of the incoming message play an essential role in determining how intelligible an utterance is. She notes three factors that contribute to speech perception: the sound produced, the setting, and the listener's previous knowledge regarding the language, topic and context. Fraser states that 'top-down' information (previous knowledge and context-based predictions) and 'bottom-up' information (acoustic information in the sound signal itself) are determinant mechanisms to understand utterances; that is, they are interrelated clues that help listeners to interpret a message. In line with this, Fraser (2018) claims that contextual knowledge and minimal acoustic cues are sufficient elements for a person to understand a word. She points out that the number of times



a person hears an utterance is irrelevant if there are not enough contextual and acoustic elements available. Thus, both the context and content of words can influence how they are understood. Everyday listeners unconsciously rely on external (background) and internal (linguistic meaning) factors for spoken-word recognition (Fraser, 2013; 2018; 2019).

The present study was focused on assessing intelligibility of the Spanish speakers' production of the vowel contrasts by means of a word recognition task; words in isolation without other related material preceding or following were chosen, so no context 'top-down' information that could help out listeners was present; avoidance of using connected texts (conversations, picture descriptions, etc.) enables the rater to focus and rely on the quality of the bottom-up acoustic performance information in the productions themselves.

### **7.1.1 Structure of the chapter**

The sections of the chapter are organised as follows:

- description of the participant raters
- method
- the set-up of the perception test
- analysis and results.

## **7.2 Perception test by Native British English Speakers**

### **7.2.1 Participants**

The participants were recruited based on their being native speakers of British English with no hearing impairment. They were contacted and invited to participate through email and social media; a voluntary response sample of 148 subjects agreed to participate.

The majority of the respondents (94%) did not have a degree in a foreign language or linguistics, no one reported speaking Spanish as a foreign language, and most of them had a bachelor's degree (see Appendix H for a full description).

### **7.2.2 Sampling summary**

The potential initial sample consisted of a total of 148 respondents. Of these, three individuals were excluded: one on the basis of not being a British English native speaker (South African); a further two were eliminated for giving incomplete responses.

The remaining 145 participants reflected a diverse range of occupations and professions.

## **7.3 Data management**

The perception test was carried out via an online survey created and hosted on the platform SoSci Survey (Leiner, 2019). Prior to data collection, and in line with the University of York and the Language and Linguistic Science Department's ethics guidelines, participants were informed about the objectives of the research (see Appendix I) and signed an electronic consent form which explicitly allowed them to withdraw at any point during data collection (see Appendix J).

All the data logged were anonymous; participants were never asked during the online survey to reveal their names. After the completion of the perception task, the data were safely stored in the SoSci Survey programme accessible only by password, then downloaded and held on a University of York secure server and backed up onto a password-protected 1TB external hard drive.

## **7.4 Method**

The main objective of the word recognition task was to determine to what extent the degree of separation between vowels produced by the Spanish speaking participants was associated with the correct identification scores obtained from the native speaker raters. Measures were obtained by having native British English speakers listen to utterances as part of an online intelligibility test. To avoid prediction during the test, isolated words were presented without top-down information to influence acoustic signal interpretations.

The following sections explain the design of the online test and the stimuli used, followed by the procedures of data collection and the conclusion.

### **7.4.1 Perception test**

#### **7.4.2 Test design**

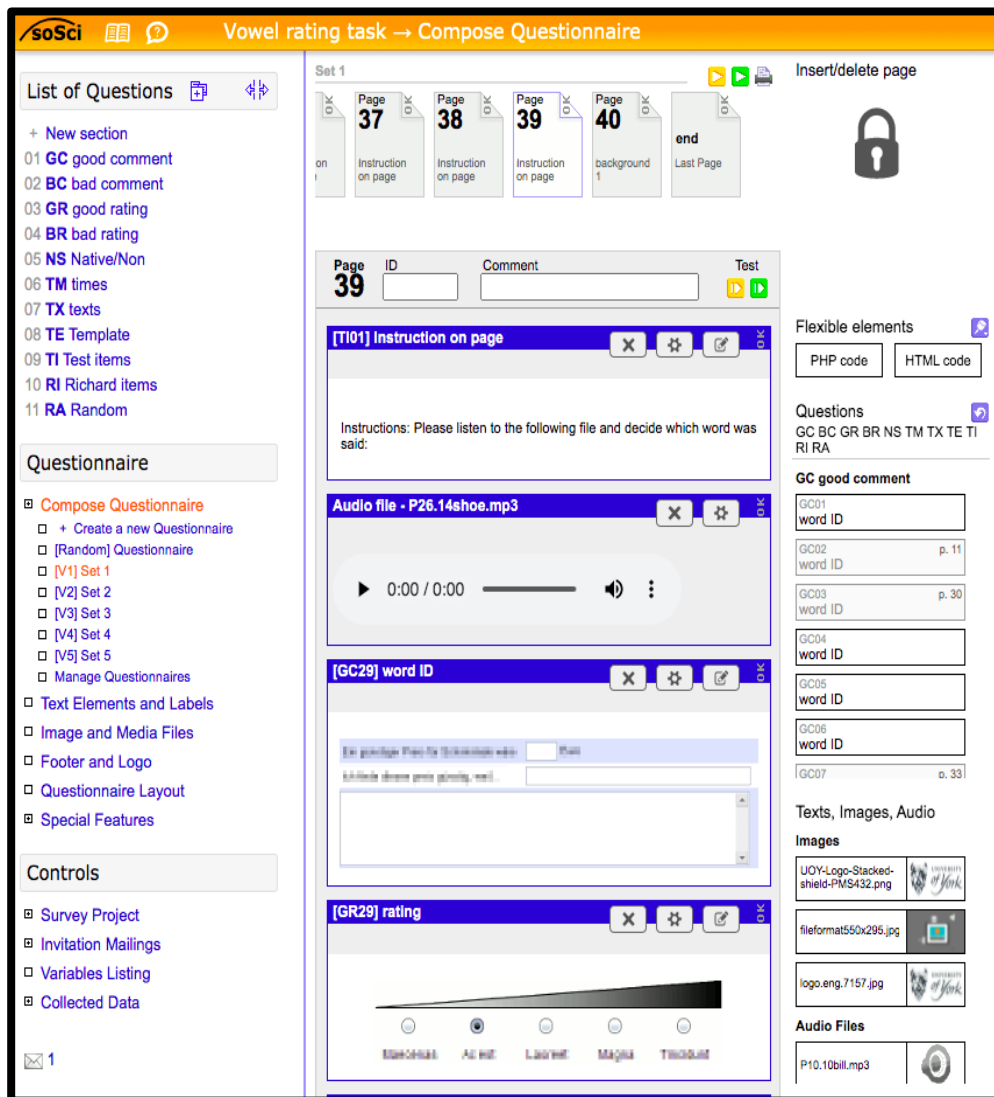
For this task an online test was designed using the programme SoSci Survey (Leiner, 2019). This platform was selected because it allows the incorporation of hundreds of audio stimuli, randomises them, and keeps track of the arbitrary sequences sent to participants in order to control the rate of answers. These features were especially useful because of the need to work with 155 different utterances and to guarantee that all of them were heard and evaluated by the participants. The final design featured five different sets of tests with one audio file per word production (35 audio files each). These five tests were randomly assigned to participants at the moment they opened the link.

Prior to using the programme, permission was requested and obtained from the SoSci Survey administrator, which allowed the free use of the programme and the option of working with audio files.

The following figure illustrates the design process.

**Figure 37**

*Picture that shows the design process of the perception test*



(<https://www.soscisurvey.de/admin/index.php>) software in the public domain.

### 7.4.3 Stimuli

The stimuli used in this test were composed of 15 isolated words of the form CVC previously recorded by ten Spanish participants selected from the group of 40 speakers. The criteria for choosing these ten speakers were based on their production of vowel distances for /i:/ and /ɪ/, /ɪ/ and /e/ and /u:/ and /ʊ/ by T3, including the five speakers with the highest separation (clear distinction), and the five speakers with the lowest separation between the vowel pairs. Contrast in duration was not achieved by any of the speakers, as stated in the results of Chapter 6, and it was therefore ignored in the selection of material for this test.

It is noted that participants belonging to the high separator group developed a qualifying degree of separation between the vowel pairs examined after T1.

A total of 150 words (3 words x 5 vowels x 10 speakers) were used. The countries represented by the speakers were Spain (6) and Mexico (4). Moreover, in order to determine that native British English participants of this perception test were able to do the task if they heard native data, five words produced by a native British English speaker with a relatively standard accent were also included in the test (as a baseline check). Overall, a total of 155 isolated words to explicitly evaluate acoustic signal interpretations without top-down information were presented as stimuli.

#### **7.4.4 Supporting information**

In addition to the audio stimuli, questions focusing on personal acuity and background information were asked, to give an indication of the level of difficulty participants had when performing the task. Their linguistic background and personal information were collected to have an overview of the participants.

##### **1) Self-perception:**

Three questions were asked after the completion of each audio task. These enquiries were related to the level of confidence in recognising the word heard, recognition of a native or non-native English speaker's voice, and finally how many times the word was heard before typing a response.


##### **2) Background information:**

The last part of the test featured nine questions that gathered information about the listener's age, gender, level of education, occupation, accent, foreign languages spoken, possible linguistic degree or background, and whether they could guess the nationality, or first language of the people heard in the recordings. In the end, this data was not systematically or extensively analysed, as this was a test of general comprehensibility of pronunciations of the Spanish learners, rather than of variability across the native English evaluators.

The following Figures illustrate how the test was visualised on a computer screen.

**Figure 38**

*Sample of the final format of the online test*

 UNIVERSITY  
*of York*

8% completed

**Instructions:**  
Please listen to the following file and decide which word was said:

▶ 0:00 / 0:00 — 🔊 ⋮

Type the word that you heard (only one word)

Rate the level of confidence in recognising the word you have heard

Not at all confident     slightly confident     fairly confident     very confident     completely confident

Based on the word you heard, would you identify the speaker as :

a native English speaker

a non-native English speaker

How many times did you listen to the word?

(<https://www.soscisurvey.de/admin/index.php>) software in the public domain.

**Figure 39**

*Sample of the final online format of for the background information questions*

The image shows a screenshot of an online survey form from the University of York. At the top left is the University of York crest and logo. To the right of the logo, the text "UNIVERSITY of York" is displayed. A progress bar in the top right corner indicates "98% completed". Below the header, the instruction "Complete the following background information :" is followed by several questions:

- What is your age?** A dropdown menu with "[Please choose]" selected.
- What gender do you identify as?** A dropdown menu with "[Please choose]" selected.
- What is the highest degree level of education you have completed?** A dropdown menu with "[Please choose]" selected.
- What is your occupation?** A text input field.
- How would you describe your English accent?** A text input field.
- Do you speak another language? If yes, which one(s):** A text input field.
- Do you have a degree in foreign languages?** A text input field.
- Do you have a degree in linguistics?** A text input field.
- Could you recognise the origin, nationality or the first language of the speaker(s) you listened in the task? If yes, write it down (e.g continent, country, region...) If not, write "no"** A text input field.

(<https://www.soscisurvey.de/admin/index.php>) software in the public domain.

#### 7.4.5 Testing Procedure

Before launching the test with participants, a pilot test was carried out to ensure clarity of the instructions, revise formal aspects such as spelling, grammar etc., check the approximate time required to complete the test, and ensure that the audio ran without glitches. After the responses of five native English-speaker participants, the test was reviewed to enhance its quality.

The final test was available to participants through a link created by the SoSci Survey (Leiner, 2019) programme. This link randomly assigned participants one specific set of tests, ensuring that each test was completed by about the same number of people.

To avoid rater fatigue and familiarity with the voices presented, each participant rated 35 words altogether, where 30 were produced by non-native speakers (mixed participants) and five were from a native British English speaker, to test consistency in the identification. Prior to the starting point of the test, raters were asked to complete the task in one sitting, using a computer in a quiet room and with headphones to minimise external noises or distractions. The test was self-paced, and it was possible to replay the audio as many times as needed.

After signing the consent form, participants were asked to familiarise themselves with the task by having a practice round with speech samples not related to the data set. Throughout the test, participants were instructed to listen and identify the words presented by typing the word they heard - following the Munro & Derwing (1995) method to assess intelligibility - to later complete additional questions related to their answers and personal background.

The test lasted approximately 15 to 25 minutes.

Thus far, the method and process to construct the word recognition task have been explained. All the content has provided potentially important data about native English speakers' perception of non-native speech, usable to support the main objectives of this research. Even though a range of data was collected, for the purpose of this study only word identification and some personal background information were chosen for the following analysis. The remaining data are kept for future analysis.



## 7.5 Data analysis

The data obtained from native British English speakers (NE) were analysed in terms of the percentage of words both correctly identified and written by the respondents (143 listeners x 15 tokens = 2,145 responses). The accurate identification of the vowels: /i:/, /ɪ/, /e/, /u:/ and /ʊ/, per participant, was then averaged to obtain the median outcome (for each vowel) for the two groups.

## 7.6 Results

The results are presented in two groups: a) higher separator and b) lower separator. Overall, the results show a correlation between the performances of the groups and the correct identifications by the native English speakers.

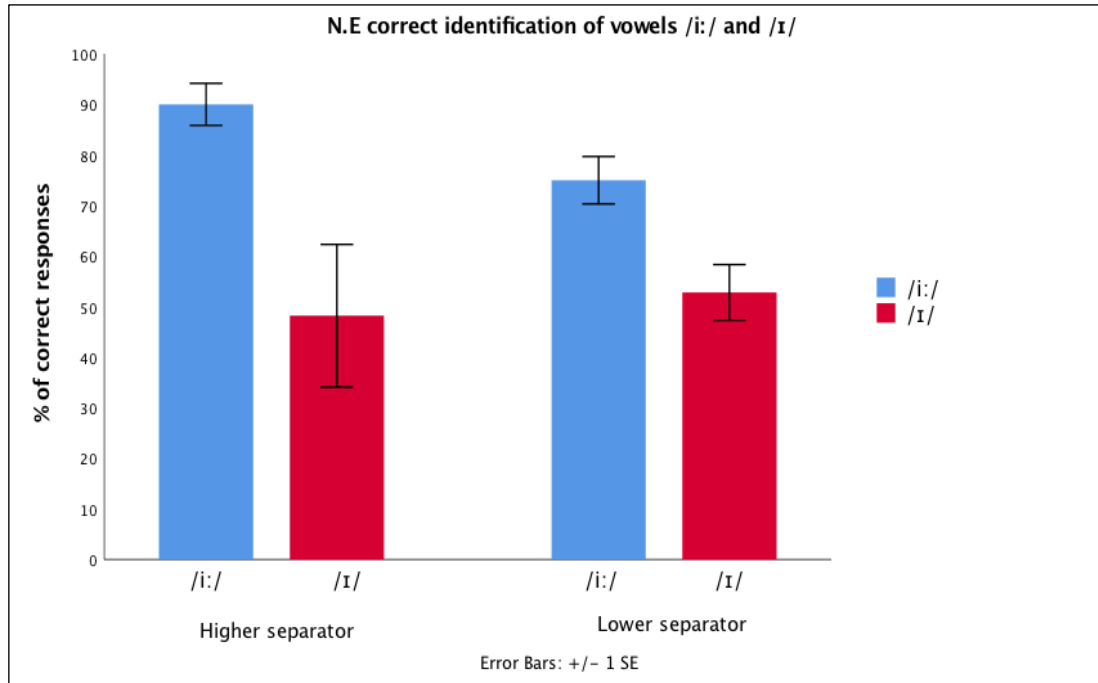
### 7.6.1 /i:/ and /ɪ/ results

- **Higher separator group:** the five sampled speakers produced a clear distinction between the pair by T2 and maintained it throughout T3. The rate of correct responses from native English speakers to /i:/ vs /ɪ/ production was 90% and 48% respectively. The /i:/ was most recognised with a tendency to be perceived as a distinctive vowel. However, the /ɪ/ showed marked confusion in discernment, where /ɪ/ was heard as /i:/ 17% of the time and as /e/ 25% of the time. The other 10% was divided into /ei/ 5% and as a non-identifiable 5%.
- **Lower separator group:** by T3, the five sampled participants did not produce a such a clear distinction between the vowel pair. The rate of correct responses from the native English speakers to /i:/ vs /ɪ/ production was 75% and 52% respectively. Overall, the vowel /i:/ was generally perceived as correct; however, a misperception was found for the vowel /ɪ/, which was interpreted as /i:/ 30% of the time, as /e/ 6%, and as non-identifiable 12% of the time.

The following figure exemplifies the results for /i:/ vs /ɪ/.

**Figure 40**

*Overall correct identification for /i:/ - /ɪ/ by N.E. speakers*



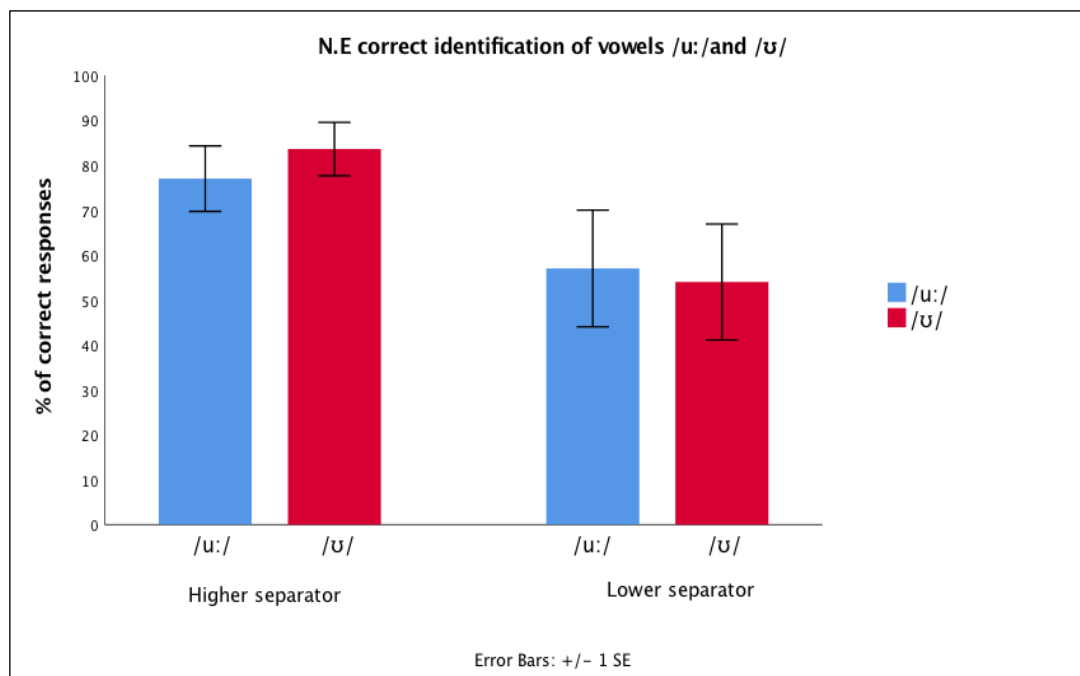
### 7.6.2 /u:/ and /ʊ/ results

- **Higher separator group:** By T3, the five sampled speakers produced a clear distinction between the vowel pair. The rate of correct responses from native English speakers to the vowel /u:/ was 77%, and for the vowel /ʊ/ 83%. The raters' accurate identification of words presenting the vowel show speakers produced them as distinctive.
- **Lower separator group:** by T3 the five sample speakers did not produce a clear distinction between the pair. The rate of correct responses from native English speakers to /u:/ vs /ʊ/ production was 57% and 54%, respectively. A misidentification pattern was found for the vowel /u:/ which was heard as the vowel /ʌ/ 16% of the time, as /əʊ/ 12% of the time, as /ʊ/ 5% and as a non-identifiable 10%. In addition, /ʊ/ was confused with the vowel /æ/ 20% of the time with /ɒ/ 10% of the time and with /ʌ/ 6%, and as non-identifiable 10% of the time.

The following figure exemplifies the results for /u:/ vs /ʊ/ vowels.

**Figure 41**

*Overall correct identification for /u:/- /ʊ/ by N.E. speakers*



### 7.6.3 /ɪ/ and /e/ results

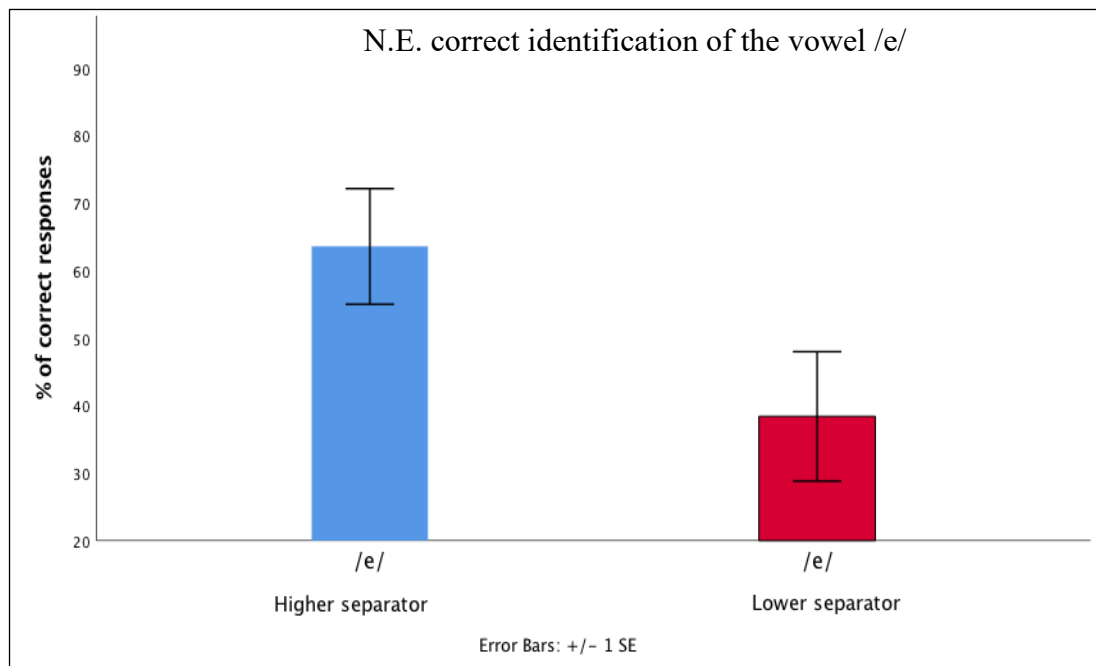
- **Higher separator group:** by T3, the five sampled speakers produced a clear distinction between the pair. The rate of correct responses from native English speakers to the /e/ vowel was 63%, whereas 37% of the respondents confused the /e/ for the vowel /æ/.

The results for /ɪ/ are the same presented in the above section a.

- **Lower separator group:** by T3, the five members of this group did not produce a clear distinction between the pair. The rate of correct responses from native English speakers to the production of the vowel /e/ was 31%. A misidentification trend was found for /e/, which was heard as /æ/ 60% of the time and as a non-identifiable 9% of the time. The results for /ɪ/ are the same as those presented in the above Section.

**Figure 42**

*Overall correct identification for /e/ by N.E. speakers*



Overall, the above results show a positive correlation between the groups' performances as measured in terms of E. distances between F1 and F2 and the intelligibility outcomes. The 'higher separator group', which presented a near- to- RP norm distinction between the vowel pairs obtained the highest ratings. In contrast, the 'lower separator group', who exhibited smaller quality distances between the members of the vowel pairs, produced vowels which were harder to discriminate by native English speakers who showed a tendency toward inaccurate segment identification.

The results support the claim that spectral acoustic cues are the primary source for identifying vowel differences for native English speakers while temporal cues play a secondary role (Hillenbrand et al., 2000; Holt & Lotto, 2006). This can be reinforced by the fact that duration differences were not present during the test. Therefore, the clear identification patterns shown by the native English-speakers raters were based on vowel quality.

## 7.7 Summary

The aim of this chapter was to determine whether the spectral changes made by the Spanish participants had consequences on how the production of vowels were heard by native British English speakers. In addition, the work aimed to corroborate the importance of acoustic cues (for native English speakers) in identifying vowels. I return to the interpretation of these results in the Discussion Chapter 10.

The following Table summarises the results for each group.

**Table 27**

Confusion matrix for vowels. The row represents the target vowels. The column represents the listeners' responses.

<b>Higher separator</b>		Perceived	/i:/	/ɪ/	/u:/	/ʊ/	/e/	/ei/	/ʌ/	/ɔ:/	/ɒ/	/æ/	/əʊ/	n/a
Target	Perceived													
/i:/		90%	10%											
/ɪ/		17%	48%			25%	5%							5%
/u:/				77%	15%			4%	4%					
/ʊ/					83%				5%		12%			
/e/						63%						37%		
<b>Lower separator</b>														
/i:/		75%	20%			5%								
/ɪ/		30%	52%			6%								12%
/u:/				57%	5%			16%				12%	10%	
/ʊ/					54%			6%		10%	20%		10%	
/e/						31%					60%		9%	

## Chapter 8. Marking of English Vowel Contrasts over a Year: Individual Performance in Connected Speech

### 8.1 Introduction

The previous chapters have presented analyses of the learners' progression over the course of a year based on analyses of word list data. The rationale for this, as argued on Chapter 5, is that word lists allow the speaker to focus most heavily on the words being pronounced and, by minimising contextual factors, give the most direct window to their phonological competence. Phonetically balanced<sup>10</sup> passages are another widely used elicitation technique. While still being a highly controlled method they offer a greater degree of ecological validity (see Lammert et al., 2020) than the word list technique and were therefore incorporated into the present research for this reason.

It was considered possible that the participants in this study would vary their vowel productions under the two elicitation conditions. Previous research on this issue has shown variable results, and the general picture is somewhat unclear. For example, in respect of vowels, Huber et al., (1999) found that F1 values tend to decrease in connected speech and increase in a clear and more articulated style, while Ferguson and Quené (2014) argue that F1 values, especially for high vowels, do not differ according to speech style. Koffi and Krause (2020) support the previous research stating that variation in speaking styles does not have an impact on formants frequencies and intelligibility is not affected. Smiljanić and Bradlow (2008) tested the length distinction between short/long vowels in different speaking styles, with results showing that vowel duration was similar regardless of the speaking context. While contributing generally to understanding of performance differences between citation form and continuous speech, i.e., reading passages, this chapter assesses the progress of the participants in developing contrasts specifically in the English vowels /i:/ and /ɪ/, /ɪ/ and /e/ and /u:/ and /ʊ/ with respect to both spectral (vowel quality) and temporal (duration) properties.

For convenience to the reader, some tables and figures to illustrate the different trajectories for individual speakers over time have been assigned to Appendices K and L.

---

<sup>10</sup> The meaning of 'phonetically-balanced' (see Chapter 4) refers to the passages containing all the phonemes of the language concerned in proportions that reflect their general incidences of occurrence (Gibbon et al., 1997).

### 8.1.1 Structure of the chapter

This chapter is organised into the following sections:

- the stimuli employed to collect data
- the analyses performed on the selected phonemes
- the results for both spectral and temporal features
- summary

## 8.2 Stimuli

The participants were recorded<sup>11</sup> reading two phonetically balanced passages at three different times during the academic year. The first task was recorded one month after arriving at University; the second task was recorded after five months of living and studying in the UK; and the final task was completed at the end of the academic year.

In each session, participants were recorded reading: *The North Wind and the Sun* (International Phonetic Association, 1999), *The Grandfather Passage* (Van Riper, 1963), *The Story of Arthur the Rat* (Abercrombie, 1964), *The Boy who cried Wolf* (adapted from Aesop's Fables), *The Caterpillar* (Patel et al., 2013), and *The Rainbow Passage* (Fairbanks, 1960). The readings were produced at a normal rate and completed within five minutes approximately, due to speakers' English proficiency level (intermediate) and the high-frequency words used in the passages.

## 8.3 Analysis

The analyses presented below were conducted with reference to both spectral (vowel quality) and temporal (duration) features. Formants analyses are presented first, followed by the distance calculated between vowel (E. distance) and the duration measurements (as stated in Chapter 5, the /e/ was excluded for the duration comparison analysis).

Only one passage was selected from each session for the analysis: Time 1 - *The North Wind and the Sun*, Time 2 - *The Story of Arthur the Rat* and Time 3 - *The Boy who cried Wolf*

---

<sup>11</sup> Participants and recording procedures are the same as described in Chapter 4

(adapted from Aesop's fable). The *Caterpillar* and *Rainbow Passage* data are kept for a forthcoming investigation.

Different reading passages were selected for each time point to avoid any possible learning effects regarding the passages themselves.

### **8.3.1 Formant analysis**

Monosyllabic words and polysyllabic words where primary stress fell on the target vowels /i:/, /ɪ/, /u:/, /ʊ/ and /e/ were selected from the reading passages across the three recordings.

A total of 2,018 tokens from the 40 speakers included in the corpus was obtained and analysed (see Table 28). /u:/ and /ʊ/ syllables with onset /j/ were not included in order to avoid fronting as a co-articulatory effect, as opposed to an adjustment to newly emerged L2 fronting norms (see Chapter 6). Likewise, no syllables with initial /w/ were included so as to avoid co-articulatory F2 reduction effects arising from the velarity and the lip rounding which also serves to reduce the frequencies of F1 and F2, and consequently make vowels sound closer and backer.

Only correctly pronounced words were considered to analyse the target vowel, i.e., productions where the speaker had clearly misidentified the word were excluded. The number of tokens per vowel fell within a range of 2-5 for each speaker. Table 28 details the token counts for each time recorded.



**Table 28***Total number of tokens per vowel across a series of time points*

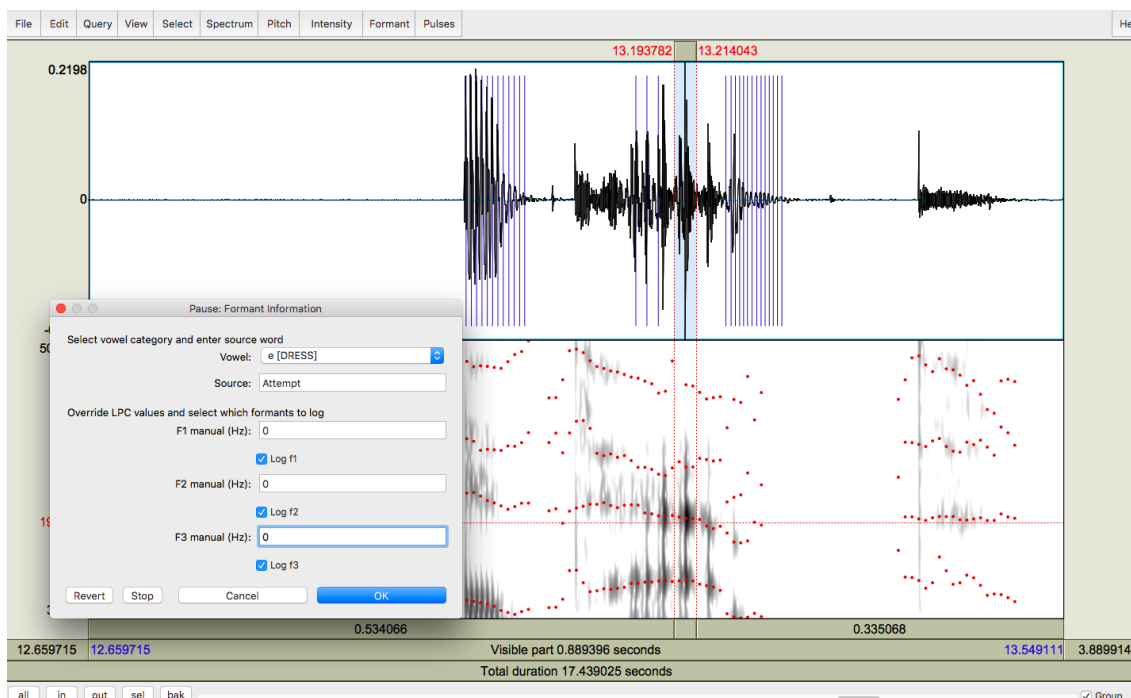
	<b>FLEECE</b>	<b>KIT</b>	<b>GOOSE</b>	<b>FOOT</b>	<b>DRESS</b>
<b>Time 1</b>	142	130	73	72	132
<b>Time 2</b>	136	214	107	100	153
<b>Time 3</b>	158	130	151	142	178
<b>Total: 2,018</b>	436	474	331	314	463

The vowel quality analyses were, as with the word list data, based on formant frequencies (F1/F2). To obtain the values, the words from the passages were first isolated using version 3.0.2 of Audacity recording and editing software (2021). The words were saved as .wav files. As previously, formant measures were manually extracted at approximately each vowel's middle point using Praat software version 6.0.49 (Boersma & Weenink, 2023) with max value settings of 5500 Hz and 4.5 as the number of formants (with adjustment when needed). Finally, the raw values of F1 and F2 were normalised using Lobanov's method in the NORM vowel normalisation software (Thomas & Kendall, 2007) to obtain the speaker mean per vowel.

The following Figure is an example of a word extracted from a reading passage where F1 and F2 were measured with Praat.

**Figure 43**

*Formant analysis sample.* The speech uttered in this extract is the word ‘attempt’. The menu on the left is used to log the formants, vowel category and source



### 8.3.2 Vowel distance analysis (F1-F2)

The degree of separation between the pair of tokens examined was again determined calculating E. distances  $d = \sqrt{(F1' - F1'')^2 + (F2' - F2'')^2}$  where F1 and F2 were considered alongside each other. These analyses were conducted for each speaker with the normalised mean value for each pair of vowels: /i:/ and /ɪ/, /ɪ/ and /e/ and /u:/ and /ʊ/.

The larger the E. distance the greater the separation between the contrast vowels produced by each individual speaker across three time points.

### 8.3.3 Vowel duration analysis

The durations of the two sets of vowel contrasts /i:/ and /ɪ/, /u:/ and /ʊ/ (1,555 tokens in total, 3 times) was measured in seconds using a Praat script hand-corrected and finally, converted into milliseconds (ms). Initial segmentation (in the form of Text Grids) was manually created

to facilitate the extraction of the duration values. The vowel measurements<sup>12</sup> were taken from the onset of voicing in the vowel to the offset of each vowel in order to be able to later analyse the speaker's mean duration per vowel. The analysis was measured only by looking at vowel durations in vowels occurring in syllables followed or closed by a voiced consonant (as in Chapters 5 and 6).

## 8.4 Results

This section presents the results obtained for formant frequencies (F1, F2) and durations separately. The outcomes include the differential rates of progress across members of the group regarding the separation of the vowels in terms of formants (E. distance) followed by the results of vowel duration differences between the two pairs of vowels (/i:/ and /ɪ/, /u:/ and /ʊ/). High and low performing individuals are identified by reference to their ascending or falling values with respect to RP normative values stated in Chapter 5.

### 8.4.1 Vowel separation results

The following results present the degree of separation between the vowel contrasts /i:/ and /ɪ/, /ɪ/ and /e/ and /u/ and /ʊ/ obtained after calculating the E. distance. A division in the results was made to estimate the level of change throughout time and the distribution/spread of the vowel pairs for all 40 speakers. The arbitrary partition was based on the E. distance reference data points for RP speakers<sup>13</sup>. For the /i:/ and /ɪ/ vowels, the dividing line was 250 Hz; for the /u:/ and /ʊ/ vowels, the value was set at 110 Hz; and for the /ɪ/ and /e/ vowels the reference threshold was 250 Hz.

In order to look at differences across time at the individual level, three main groups were established to identify speakers who produced changes in formant values (in terms of separation of vowels) and individuals who did not. The groups were named: (a) Moderate/static (b) Substantial/large movement and (c) Backward movement

---

<sup>12</sup> The length distinction for tense vowels, in this case, /i:/ and /u:/ was considered by examining syllables closed by a voiced consonant only (see previous Chapter 5 for explanation).

<sup>13</sup> The values were taken from Deterding (1990 as cited in Deterding, 1997) and used to calculate the E distance between the values examined, see Chapter 6.

a) *Moderate/static* group: this group was composed of participants who did not produce changes in formant values (vowel separation) above the thresholds established for each vowel pair (e.g., results below 250 Hz for /i:/ and /ɪ/, below 110 Hz for /u:/and /ʊ/, and below 250 for /ɪ/ and /e/).

b) *Substantial/large movement* group: participants assigned to this group were the ones who showed changes in formant values above the thresholds established for each vowel pair or maintained their formant values above the RP norms (e.g., results above 250 Hz for /i:/ and /ɪ/, above 110 Hz for /u:/and /ʊ/, and above 250 for /ɪ/ and /e/).

c) *Backward movement* group: this group consisted of participants who produced formant values (in terms of separation of vowels) above the thresholds set for the vowel pairs but at some testing point (T2 or T3) produced values below the thresholds. (see Appendix K).

#### **8.4.2 /i:/and /ɪ/ results**

This section shows the results for /i:/ and /ɪ/ vowel separation. Overall, the results show that most of the participants did not exhibit major progress in distancing /i:/ from /ɪ/. That is, the increase in percentages was made for the group which maintained or did not cross the separation threshold of 250 Hz.

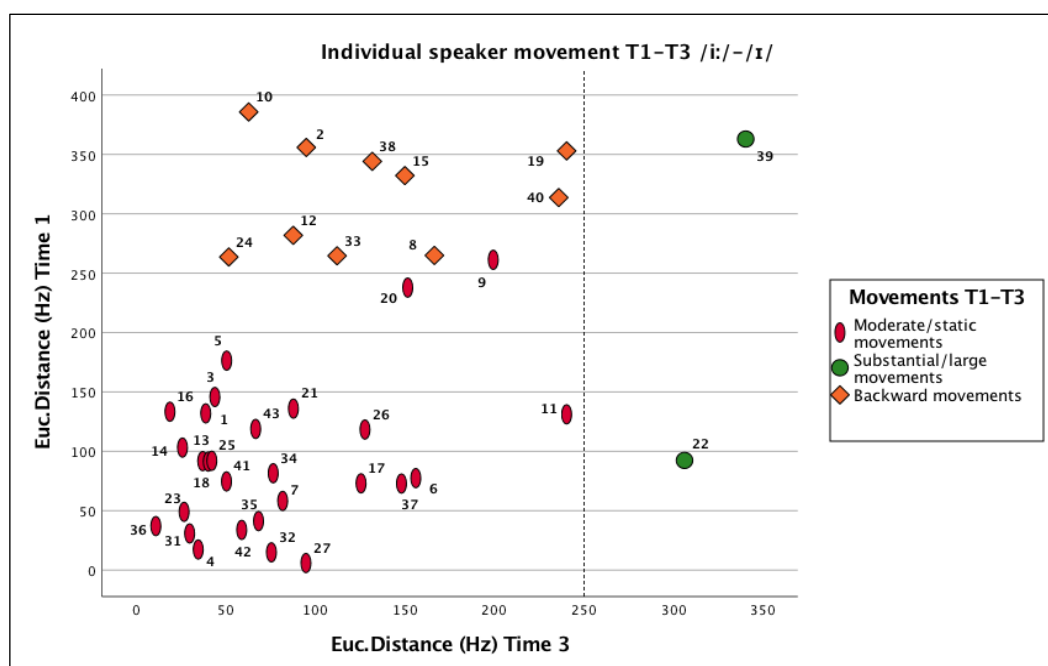
The results (from T1 to T2) show that 25 speakers maintained their separation under the 250 Hz band criterion with an average of 75 Hz, and nine speakers moved backwards to match their initial performance with an average of 150Hz. In total the 34 low performing speakers kept their separation values with an average of 94 Hz below the RP norm. However, exceptions to this trend were exhibited by six high performing speakers, which by T2 had markedly separated themselves from the group with an average distance of 296 Hz, a result that is closer to the RP value. It should be noted that three out of six speakers retained a distance above 250 Hz from T1.

Results from T2 to T3 show that 38 low performing speakers produced values with an average of 93 Hz, which falls below the British English 250 Hz set band. Only 2 high performing speakers maintained their degree of separation above the threshold from T2 to T3, with an average of 323 Hz. That is, instead of increasing the number of high performing speakers by T3 the results surprisingly reveal an opposite tendency.

The next Figure provides an overview of the results obtained from T1-T3.

**Figure 44**

*Vowel distance variations between /i:/ and /ɪ/ from Time 1-3*



**Table 29**

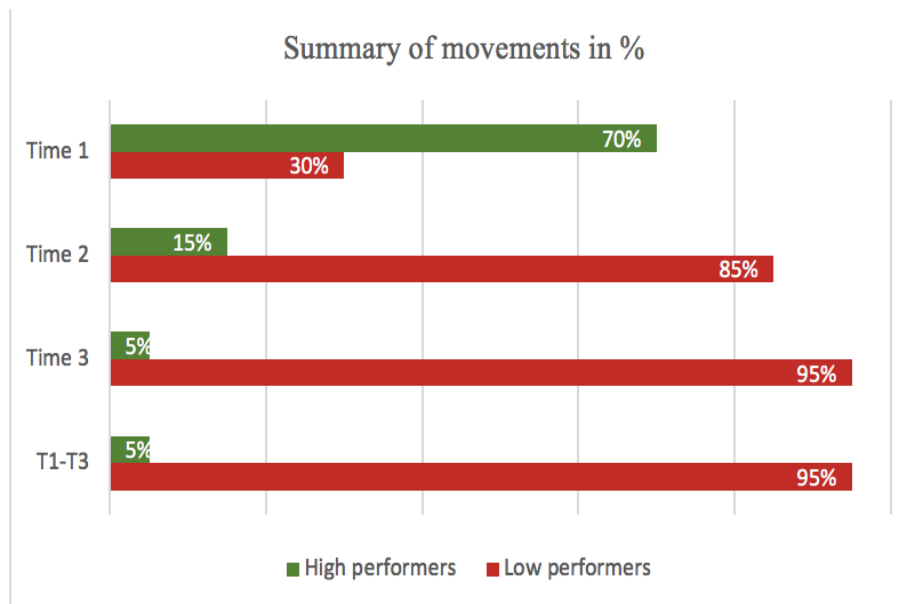
<i>Summary T1-T3</i>	<b>N</b>	<b>%</b>
a) Moderate/static movement (low performers)	28	70%
b) Substantial/large movement (high performers)	2	5%
c) Backward movement (low performers)	10	25%

Overall, in a year (T1-T3), the results indicate that 38 low performing speakers did not increase their vowel separation and maintained their values with an average of 93 Hz, which is well below the RP set value. Only 2 high performing speakers produced their vowel separations above the threshold with an average of 323 Hz, indicating a shift towards - and beyond - the British English norm.

The following Figure displays a summary of the results across Times.

## Figure 45

*Summary in percentages about the proportion of speakers who fell into either of the two categories*



In the section that follows, the results are presented in terms of F1 and F2 separately, in order to establish whether the distance between vowels was produced because participants made a distinction between F1 and F2 separately, or with both formants together.

### 8.4.3 F1 and F2 results

The results set out above were obtained based on the E. distance calculation, where F1 and F2 were considered together. Considering the outcomes, it was important to look in more detail at what may have caused the distance - or the lack of it - between the pair of vowels under examination, i.e., if there were changes, were these mainly brought about by shifts in F1 or F2? The patterns are shown below.

The majority of F1 values for both /i:/ and /ɪ/ by T2 decreased, resulting in a higher position of /i:/ (average = 337 Hz) and (at the same time) a higher location for /ɪ/ (average = 378 Hz) for most of the speakers. It was, however, observed that the number of speakers producing higher F1 values for /i:/ and /ɪ/ are comparable.

Similar changes in F2 values were found for /i:/ (average = 1803 Hz) and /ɪ/ (1699 Hz). The majority of speakers (34) reported a decreased F2 with both vowels moving towards a less fronted position by T2.

Results from T2 to T3 show a marked difference from the previous times, especially in F2. Tables 30-31 show the summarised outcomes<sup>14</sup>.

**Table 30**

*Overall results showing the number of speakers producing higher or lower F1 values from Time 2-3*

Vowel	N° sp	F1	Vowel	N° sp	F1
/i:/	38	Higher values (avg: 394 Hz)	/ɪ/	28	Higher values (avg: 415 Hz)
	2	Lower values (avg: 362 Hz)		8	Lower values (avg: 385 Hz)

**Table 31**

*Overall results showing the number of speakers producing higher or lower F2 values from Time 2-3*

Vowel	N° sp	F2	Vowel	N° sp	F2
/i:/	28	Increased values (fronted) (avg: 1894 Hz)	/ɪ/	36	Increased values (fronted) (avg: 1817 Hz)
	10	Decreased values (central) (avg: 1823 Hz)		1	Decreased values (central-back) (avg: 1787 Hz)

<sup>14</sup> I refer to a central position for /i:/ when there is a backward movement in F2 (with consideration to its position as a front vowel). And I refer to the central-back position for /ɪ/ when there is a backward movement in F2 (with consideration to its position as a front vowel).

As shown in Table 30, F1 values increased by T3 for both /i:/ and /ɪ/. As a result, both vowels were located in lower positions compared to the previous time. Speakers who produced lower vowels by T2 continued to do so, and more participants followed this trend. In general, /i:/ lowered its position towards /ɪ/ while /ɪ/ retained its lower position.

In terms of F2, a shift occurred for most of the speakers. The rise of F2 was steady for /i:/ and /ɪ/, obtaining a more fronted position compared to T2. In short, it can be said that by T3 most of the speakers produced a lower /i:/ and a fronted /ɪ/.

Overall, over a year, the changes in the values of F1 for /i:/ and /ɪ/ indicate that for /i:/ 36 speakers produced higher vowel values, which indicated that both vowels were produced in a lower acoustic location than in T1. For /ɪ/, 28 speakers increased their F1 values, indicating a lower placed vowel than in T1.

In addition, F2 results for /i:/ show a trend toward a retracted (central) vowel (within the front vowel parameters), and interestingly /ɪ/ presents a division between the number of speakers producing a more fronted vowel and the number of speakers producing a more centralised/retracted vowel, with a mean difference of 19 Hz by the end of the year. Another important observation is that by T2 the same 34 speakers who produced a central /i:/ also produced a centralised-retracted /ɪ/ (lower F2).

By T3 in terms of F1 and F2, 31 speakers who produced a lower /i:/ also produced a lower /ɪ/, and 26 speakers who produced a fronted /i:/ also produced a fronted /ɪ/ (higher F2). Finally, 22 subjects were able to produce a low and fronted /i:/ as well as a lowered and fronted /ɪ/. In summary, regarding F1, /i:/ generally tended to move towards /ɪ/, and regarding F2 /ɪ/ mostly moved towards /i:/.

#### **8.4.4 /u:/ and /ʊ/ results**

The following results present the degree of separation between /u:/ and /ʊ/ obtained by calculating the E. distance from the normalised data. The arbitrary division of the progression groups was established at 110 Hz based on the reference data points from contemporary RP previously presented.

In general, the comparison between times reflected a tendency for speakers (+70%) to noticeably maintain their initial production of vowels beyond the 110 Hz band throughout the year.



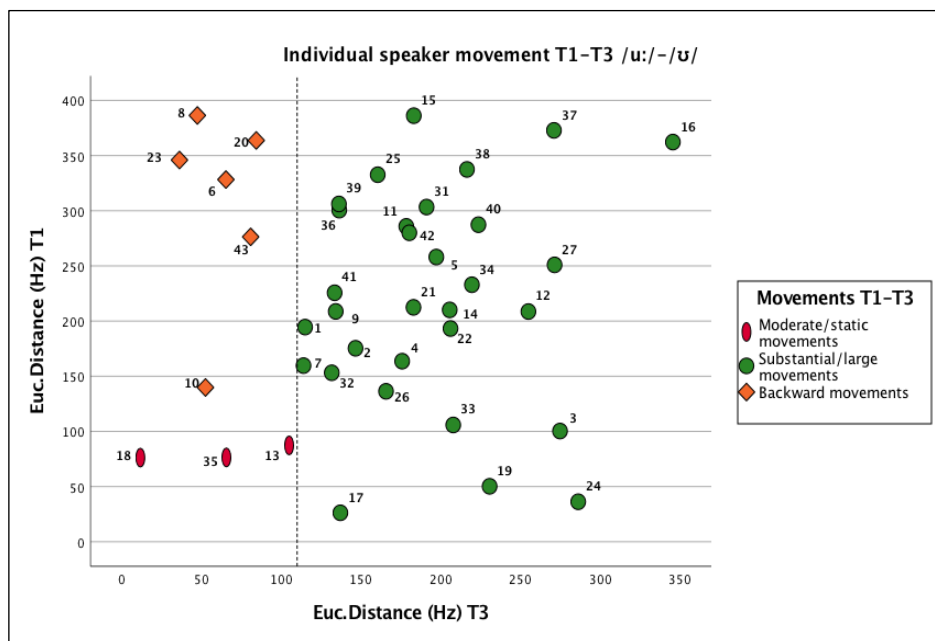
Results at T1 show that 32 high performing speakers produced distant vowels with a range value (which was kept until the end of T2) above 110 Hz. Interestingly, at T1, nine low performing respondents who produced native-like distance values (above 110Hz), by T2, reversed the position, producing less distinct vowels - below the British English norm of 110 Hz.

By T3, the results show that a considerable number of high performing speakers (32) maintained their outcomes above 110 Hz during the three times. Minimal variation in the number of performers per group was observed. By T3 eight low performers produced the vowels with an average distance of 61 Hz.

The next figure provides an overview of the results obtained by T1 and T3.

**Figure 46**

*Vowel distance variations between /u:/ and /ʊ/ from Time 1-3*



**Table 32**

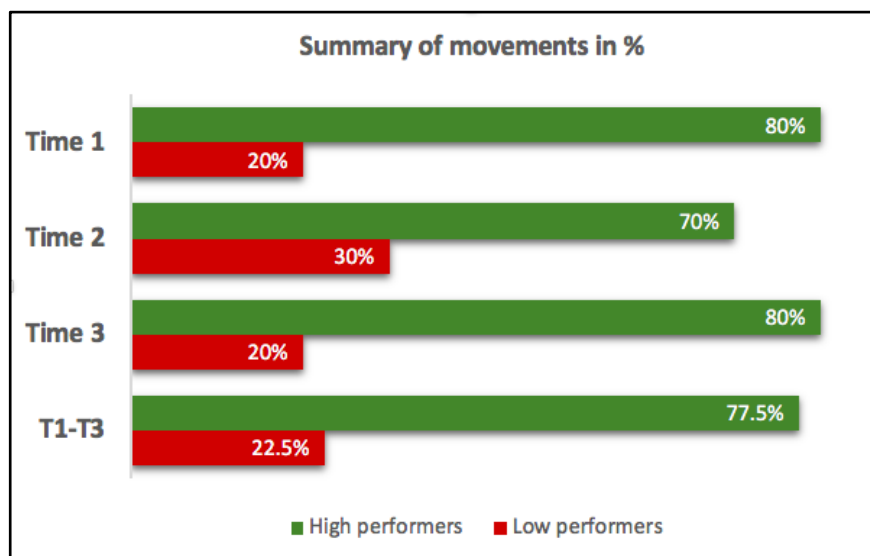
<i>Summary T1-T3</i>	N	%
a) Moderate/static movement (low performers)	3	7.5%
b) Substantial/large movement (high performers)	31	77.5%
c) Backward movement (low performers)	6	15%

Overall, in a year (T1-T3), as can be seen from Figure 46, few low performers increased their separation vowel difference above the RP set band of 110 Hz, and the high performing speakers showed a steady outcome during the three times examined, with an average distance of 193 Hz. Nine low performers produced their vowel separation below the target with an average of 61 Hz.

Figure 47 presents a summary of all outcomes at different times.

**Figure 47**

*Summary in percentages about the proportion of speakers who fell into either of the two categories*



The next section shows the results in terms of F1 and F2 separately, in order to establish whether the distance between vowels was produced because of participants' changes in F1 or F2 or both formants altogether.

**8.4.5 F1 and F2 results**

Results detailing the acoustic changes in the vowel qualities for /u:/ and /o/ are set out below. Generally, the F1 values for both vowels had a tendency to be higher when compared to participants' starting point (T1). In terms of the F2 outcomes for /u:/, the results show that most of the speakers tended to reduce their F2 values over the year, while results for the vowel /o/ show that most of the speakers increased their F2 values over the times tested.

Results show that by T2, the proportion of participants that presented F1 values for /u:/ and /ʊ/ higher than at T1 was 50:50, with an average for /u:/ of 394 Hz and /ʊ/ of 426 Hz. In terms of F2, a similar tendency was observed for /u:/ and /ʊ/ related to the proportion of participants producing lower F2 values by T2 than at T1 was 50:50, with an average for /u:/ of 1371 Hz and for /ʊ/ of 1271 Hz.

**Table 33**

*General results showing the number of speakers producing higher or lower F1 values from Time 2-3*

*Time 2-3 F1*

Vowel	N° sp	F1	Vowel	N° sp	F1
/u:/	37	Higher values (avg: 451 Hz)	/ʊ/	29	Higher values (avg: 452 Hz)
	1	Lower values (avg: 404 Hz)		9	Lower values (avg: 438 Hz)

**Table 34**

*General results showing the number of speakers producing higher or lower F2 values from Time 2-3*

*Time 2-3 F2*

Vowel	N° sp	F2	Vowel	N° sp	F2
/u:/	7	Increased (fronted) (avg:1323 Hz)	/ʊ/	34	Increased (fronted) (avg: 1401 Hz)
	31	Decreased (retracted) (avg: 1259 Hz)		5	Decreased (retracted) (avg:1447 Hz)

A clear outcome is seen in Tables 33-34. By T3 higher F1 values and the number of speakers producing them increased in apparent time for both /u:/ and /ʊ/ with a similar average for both vowels of 451 Hz. There were, however, different participant patterns for F2 for /u:/ and /ʊ/. For /ʊ/, the number of speakers producing the vowel with higher F2 values increased, and

thus, the result was a more fronted vowel compared to T2. In contrast, , F2 values for /u:/ show that most of the speakers produced an even more retracted variant than by T2.

From the results described above, it is clear that by T3, the same 27 speakers who produced higher F1 values for /u:/ also did so for /ʊ/. And the 27 speakers who produced a retracted /u:/ also produced a fronted /ʊ/.

Overall, F1 results from T1 to T3 show that the number of participants producing both /u:/ and /ʊ/ with higher values increased. The overall F1 values for /u:/ rose from an average of 343 Hz at T1 to an average of 450 Hz by T3. The general F1 results for /ʊ/ show an increase in the values from 374 Hz at T1 to 448 Hz by T3.

In terms of F2 for /u:/, the results show an increase in the number of participants producing a gradual decrease in F2 values, going from 1416 Hz at T1 to 1262 Hz by T3. That is, 37 speakers shifted towards a more retracted vowel.

The F2 results for /ʊ/ show that 23 participants gradually increased their values from 1295 Hz to 1409 Hz, producing a more fronted vowel at the end of the year.

This section covered the results obtained for /u:/ and /ʊ/ vowel distance by looking at F1 and F2 separately. In conclusion, it is possible to say that regarding F1 the vowel /u:/ generally tended to move towards /ʊ/ (higher F1), and the vowel /ʊ/ (possible due to the /u:/ ‘push’ movement) also increased its lower position within the vowel space for the majority of the speakers (33). In addition, regarding F2, the vowel /ʊ/ strikingly shifted from a retracted to a more fronted position, whereas /u:/ was gradually produced with lower F2 values over the year’s course. I refer back to these results and the comparison with the RP values in the summary section.

#### 8.4.6 /ɪ/ and /e/ results

The results for the separation between /ɪ/ and /e/ are presented below. The division of the formed groups was settled at 250 Hz (the threshold was fixed based on the reference points from the RP values). Overall, speakers exhibit fewer variations in the separation of vowels over time. 90% of the speakers produced distances greater than 250 Hz at T1. However, there was a reduction in distance by T2, and by T3, most of the speakers had increased their vowel distances above the threshold to values comparable with their initial performance.

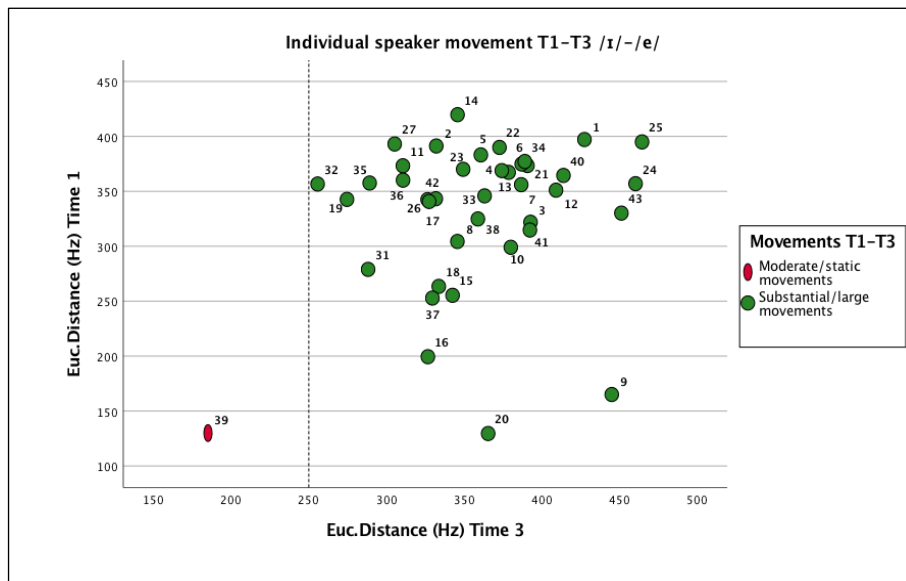
The results by T2 show that 19 low performing speakers did not produce the vowel separation /ɪ/ and /e/ above the RP set norm. Within that group, 16 participants who had initially produced values above the 250 Hz band by T1 had decreased their values below the threshold with an average of 190 Hz. In addition, 20 high performing speakers maintained their separation values above 250 Hz with a mean of 294 Hz.

By T3 an impressive improvement in the separation of both vowels was achieved by all participants (except for one). Thirty-nine high performing speakers increased their final values from an average of 246 Hz to an average of 361 Hz. Only one low performing speaker over the three-time periods did not approach the threshold level of 250 Hz separation and maintained the vowel distance with an average of 185 Hz.

Figure 48 and Table 35 show an overview of the results obtained from T1 and T3.

**Figure 48**

*Vowel distance variations between /ɪ /and /e/ from Time 1-3*



**Table 35**

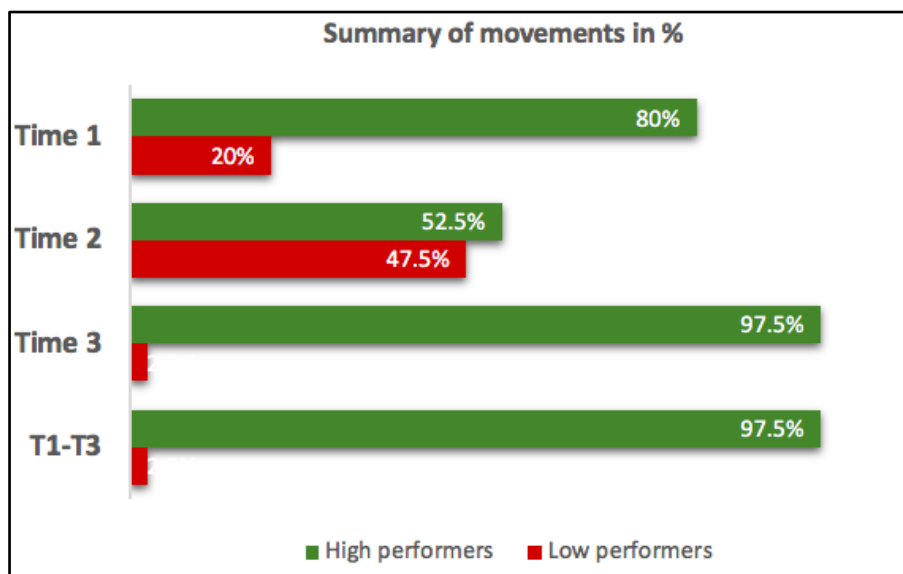
<i>Summary T1-T3</i>	<b>N</b>	<b>%</b>
a) Moderate/static movement (low performers)	1	2.5%
b) Substantial/large movement (high performers)	39	97.5%

Overall, the above results show a marked increase in the separation between /ɪ/ and /e/ between T1 (246 Hz) and T3 (361 Hz). However, between T1 and T2 40% of the speakers reduced the distance values below 250Hz.

The following Figure displays a summary of the results across all three time points.

**Figure 49**

*Summary in percentages about the proportion of speakers who fell into either of the two categories*



In the section that follows, results are presented in terms of F1 and F2 separately.

#### **8.4.7 F1- F2 Results**

Overall, over a year, it can be said that F1 values for /ɪ/ and /e/ increased. Possibly the fact that /ɪ/ came to occupy a lower position (higher F1) caused the movement (mini 'push chain') of /e/ to a lower location as well. Concerning F2, both vowels showed a tendency towards an overall increase in the frequency of the second formant, indicating production of both vowels with a more fronted tongue position than at T1<sup>15</sup>.

In terms of F1 by T2 for /ɪ/, 21 speakers produced lower values compared to T1, and 38 speakers produced lower values for /e/, going from 601 Hz to 510 Hz.

In terms of F2 and equal number of speakers produced lower values for both vowels with an average for /ɪ/ of 1699 and for /e/ of 1492 Hz.

---

<sup>15</sup> I refer to the central-back position for /ɪ/ when there are low values in F2. The same applies to /e/.

**Table 36**

Overall results showing the number of speakers producing higher or lower F1 values from Time 2-3

Vowel	N° sp	F1	Vowel	N° sp	F1
/i/	28	Higher values (avg: 415 Hz)	/e/	40	Higher values (avg: 643 Hz)
	8	Lower values (avg: 385 Hz)		0	Lower values (avg: 0)

**Table 37**

Overall results showing the number of speakers producing higher or lower F2 values from Time 2-3

Vowel	N° sp	F2	Vowel	N° sp	F2
/i/	36	Increased values (fronted) (avg: 1817 Hz)	/e/	29	Increased values (fronted) (avg: 1580 Hz)
	1	Decreased values (central-back) (avg: 1823 Hz)		10	Decreased values (central-back) (avg: 1508 Hz)

In terms of F1 and F2, the above tables show a clear pattern for both vowels.

The F1 results for /i/ and /e/ show that most of the speakers shifted their production from low values at T2 to higher values by T3. The F1 values for /i/ were raised by 28 speakers from 378 (T2) Hz to 415 Hz (T3), and the F1 values for /e/ were increased by all the speakers from 510 Hz to 643 Hz.

In terms of F2, by T3, most of the speakers (36 for /i/ and 29 for /e/) had changed the production of the vowels towards a more fronted position. The value for /i/ increased from 1712 Hz to 1817 Hz, and for /e/ from 1486 Hz to 1580 Hz.



Overall, regarding F1, 28 speakers produced /ɪ/ with higher values (lower vowel) shifting from an average of 383 Hz at T1 to 415 Hz by T3. In addition, F1 values for /e/ were increased by 37 speakers, shifting from 593 Hz to 643 Hz. Concerning F2, an equal number of speakers increased and decreased the values for both vowels.

Another point of interest is that by T3, the same 27 speakers who produced a lower and fronted /ɪ/ also did it for /e/. Additionally, 23 out of 40 speakers who by T3 produced the /ɪ/ and /e/ with higher F1 values also produced the /ɪ/ and /e/ with greater F2 values.

In summary, /ɪ/ tended to move to a lower and fronted position and /e/ was also inclined towards a lower and fronted position, possibly owing to a systemic 'push' effect, but also as an adjustment towards native speaker norms (a vowel close to cardinal 3).

The T1 to T2 data indicate a temporary reverse trend, both vowels at T2 being produced in higher and more retracted positions. I refer back to these results and the comparison with the RP values in the final summary section of this chapter.

Together these results provide important insights into the progress that participants have made in developing the contrast of English vowel pairs /i:/ and /ɪ/, /u:/ and /ʊ/ and /ɪ/ and /e/ with respect to spectral features. Also, it has been possible to observe if the progress or changes that speakers made regarding vowel contrasts have been done in terms of F1 and F2 separately or both formants together.

Despite the degree of separation between the vowel pairs examined participants generally showed some formant differentiation. That is, the overall realisation of vowels by the Spanish speakers tends to follow the British English tendency, namely: /u:/ produced in a higher position than /ʊ/, /i:/ produced in a higher position than /ɪ/, and /e/ produced lower than /ɪ/. These results and their comparison to the reference values from the contemporary RP data will also be addressed in the final summary section below.

## 8.5 /u:/ and /ʊ/ fronting results

As covered in Chapter 6, /u:/ and /ʊ/ vowel fronting has been an ongoing feature in many English varieties (see Harrington et al., 2008; Strycharczuk & Scobbie, 2017). Thus, to determine if this fronting trend was reflected in the participants from this research the fronting of back vowels was analysed in general terms.

F2 normalised values (645 in total) were considered to report the changes in high-back vowels across time. Based on the results, three groups were created to signal the formant changes: *Fronting*, *Backing* and *Same* (no changes in F2 between times) groups. A group assignment threshold  $\pm 20$  Hz was set to distinguish progress groups.

*Fronting* group: participants assigned to this group had a fronting F2 movement of +20 Hz from Time to Time.

*Backing* group: participants assigned to this group retracted their vowels +20 Hz from Time to Time.

*Same* group: participants assigned to this group had minimal variation in F2 between Times (1-20 Hz)

The results for /u:/ and /ʊ/ fronting are presented separately in the following sections (see Appendix L).

### 8.5.1 /u:-fronting

The results show that overall realisation at T1 was 1497 Hz. By T2, 27 speakers (who initially produced a more fronted /u:/ vowel, avg: 1547) retracted their F2 values (avg: 1352 Hz). Four speakers maintained their initial F2 values (avg: 1526 Hz), and nine speakers increased their original F2 values from an average of 1417 Hz to an average of 1595 Hz.

Interestingly, by T3, the results show a striking decrease in F2 values (from T2 to T3).

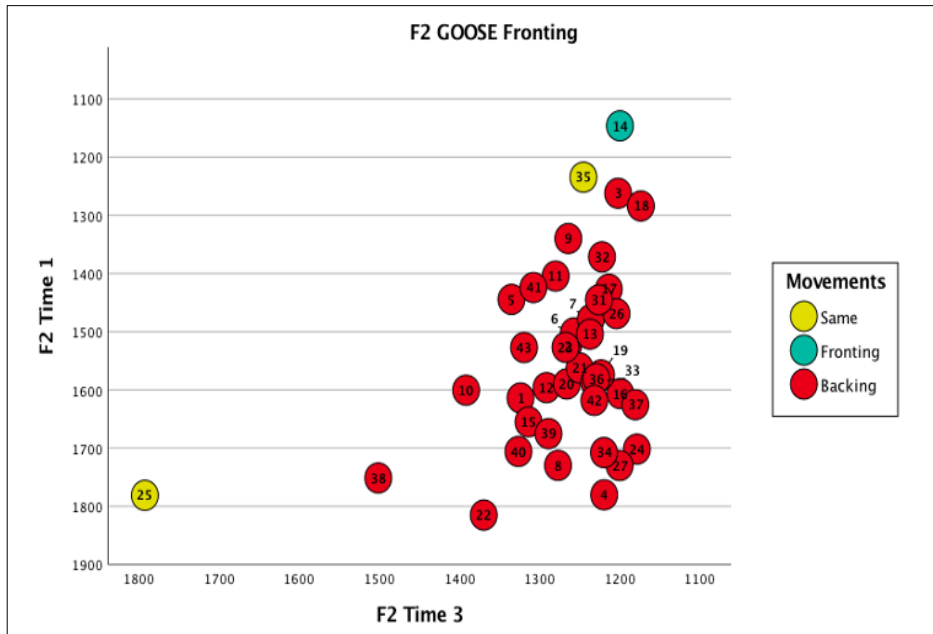
Twenty-nine speakers backed their production of /u:/ from an average of 1489 Hz to 1260 Hz.

In other words, the majority of the speakers did not show /u:/ fronting. A move towards fronting by seven speakers was seen by T2 with an average of 1313 Hz, but by the last recording session the F2 values dramatically decreased.

Figure 50 shows the overall performance from T1-T3.

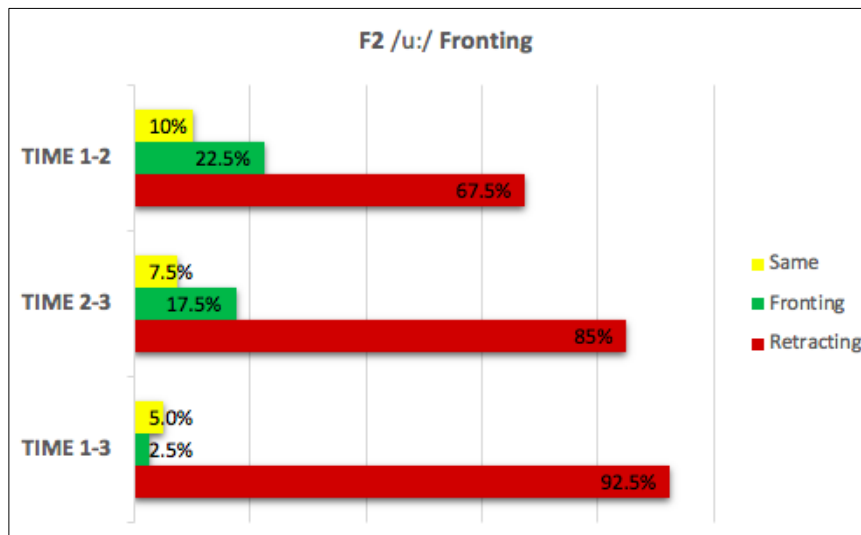
**Figure 50**

*Changes in the F2 values for /u:/ between T1 and T3*



**Figure 51**

*Summary in percentages about the proportion of speakers who fell into any of the three categories*



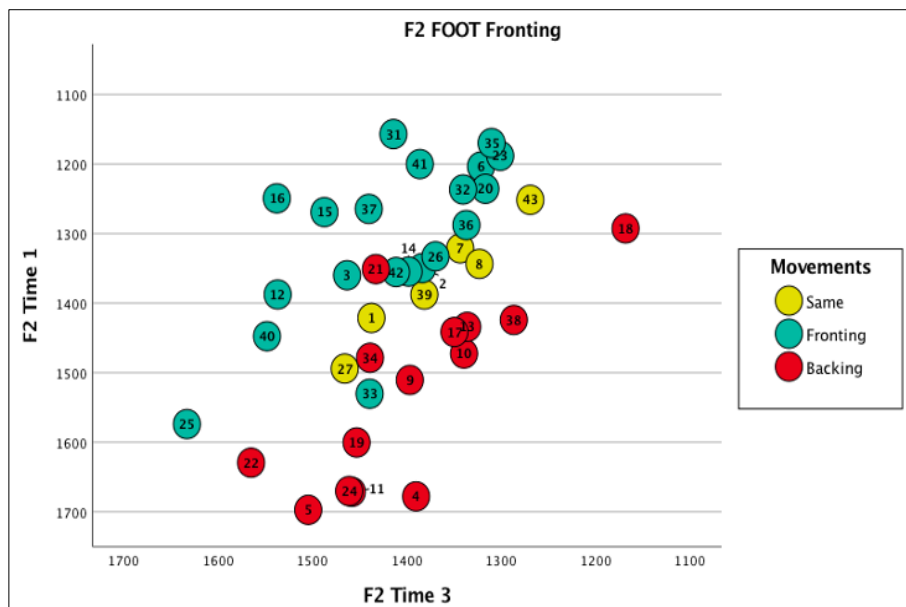
### 8.5.2 /ʊ/- fronting

The results for the production of /ʊ/ show that at T1 the vowel was realised with an average of 1393 Hz. By T2, 24 participants decreased their F2 values for /ʊ/ from an average of 1448 Hz to 1260 Hz. However, it was also observed that 11 speakers increased their formant values from an average of 1299 Hz to 1420 Hz, producing a more fronted vowel; and five speakers did not produce noticeable changes between these two times and maintained the values with an average of 1330 Hz.

By T3, 31 speakers increased their F2 values for /ʊ/ from 1267 Hz to 1409 Hz, producing a more fronted /ʊ/. It was also observable that four of the speakers produced the /ʊ/ with lower values, going from 1676 Hz to 1474 Hz, and five speakers kept their F2 values with an average of 1314 Hz. These results show that even though there was a back and forward movement among the speakers, the majority increased their F2 values and produced a more fronted vowel than at T1.

**Figure 52**

Changes in the F2 values for /ʊ/ between T1 and T3



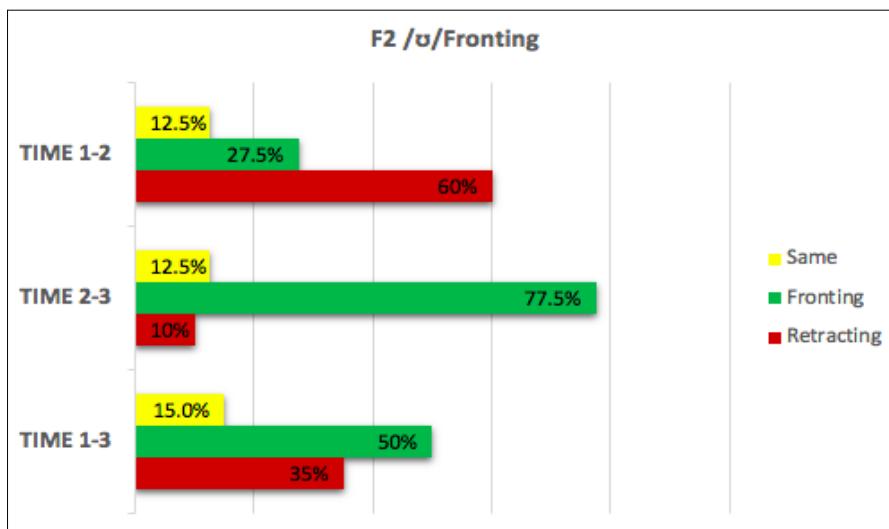
Overall, in a year, 20 participants increased their F2 values, from an average of 1307 Hz to 1418 Hz; 14 speakers produced the vowel with lower F2 values, decreasing from an average

of 1525 Hz to an average of 1398 Hz; and 6 participants did not change their values, maintaining them with an average of 1370 Hz. In general, participants' starting point for the production of F2 values was 1393 Hz average. The main variations above that value were observed especially by T3, where 31 participants individually, produced variations in the second formant producing the vowel /ʊ/ in a fronted position.

The following figure presents a general summary of the overall results.

**Figure 53**

*Summary in percentages about the proportion of speakers who fell into any of the three categories*



A summary of the formant findings is presented below, after the duration analyses in the chapter summary.

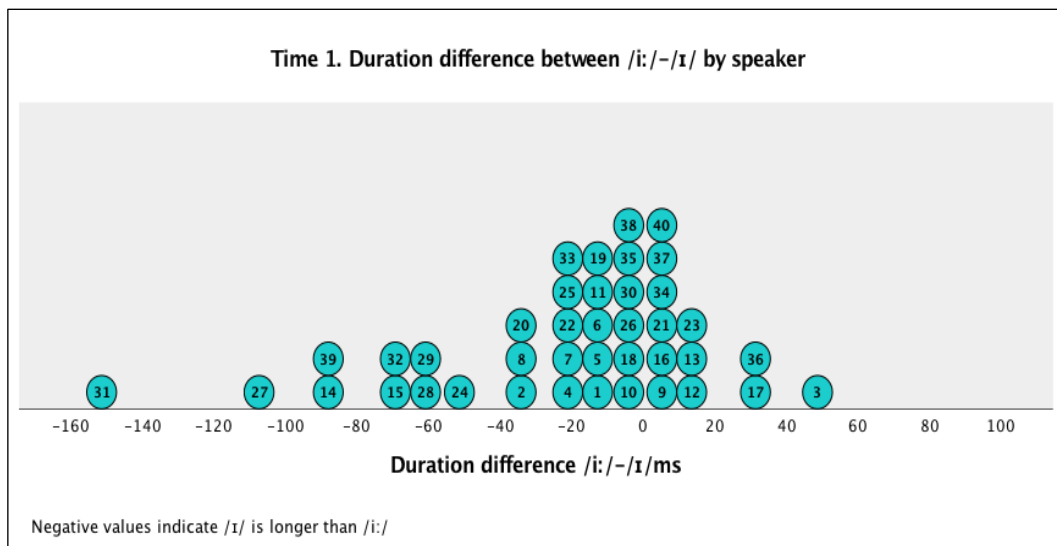
## 8.6 Vowel duration results

### 8.6.1 /i:/and /ɪ/ vowels

The quantity distinction between /i:/ and /ɪ/ vowels are presented in this section. Based on the results two groups were formed: positive values (correct length) and negative values (incorrect length<sup>16</sup>). Only vowels followed by a voiced consonant were analysed (see page 134 above).

**Figure 54**

*Duration difference between /i:/ and /ɪ/ at Time 1*



**Table 38**

*Summary T1*

<i>Summary T1</i>	N	%
a) Positive values	12	30%
b) Negative values (/ɪ/ longer than /i:/)	28	70%

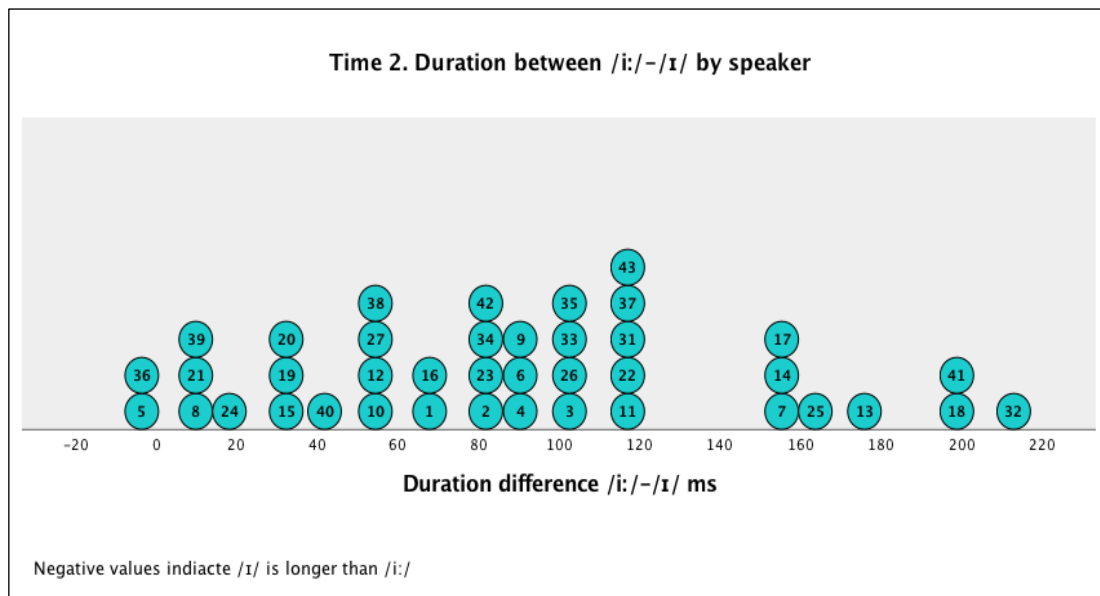
Figure 54 shows that, at T1, 28 speakers produced the vowel /ɪ/ longer than the /i:/. The average length for the production of /i:/ was 222 ms and for /ɪ/ was 264 ms. 12 speakers

<sup>16</sup> Negative values indicate that /ɪ/ was produced longer than /i:/ when it should have been shorter.

produced the correct length distinction (just by means of /i:/ being 8 ms longer than /ɪ/) with an average for /i:/ of 231 ms and /ɪ/ of 223 ms. Although the average direction of the duration of these two vowels is ‘somehow correct’ it does not approach the reference values of British English.

**Figure 55**

*Duration difference between /i:/ and /ɪ/ at Time 2*



**Table 39**

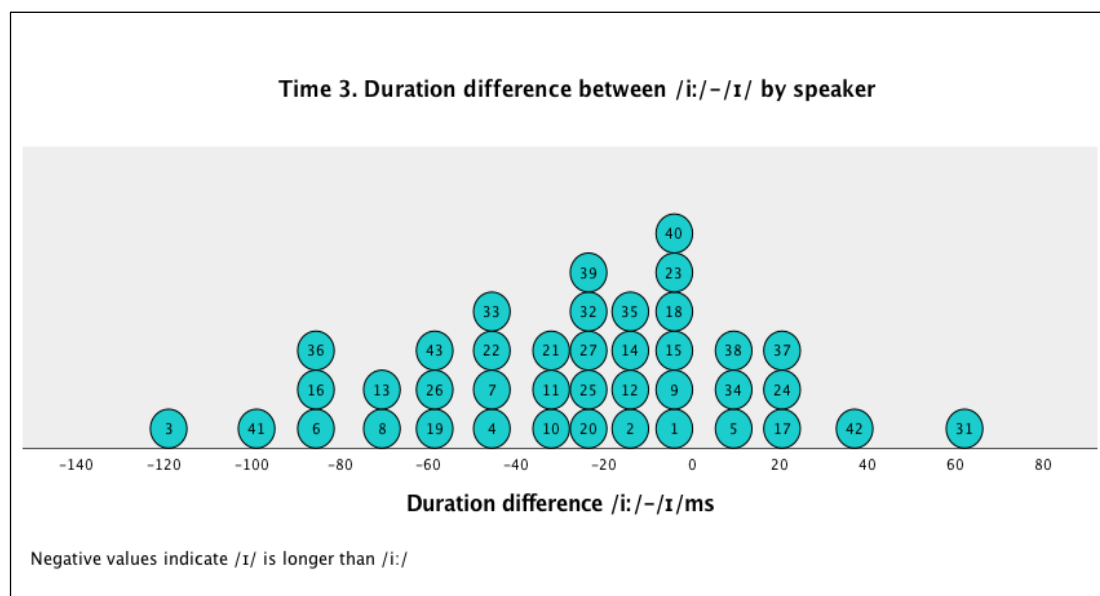
*Summary T2*

<i>Summary T2</i>	N	%
a) Positive values	38	95%
b) Negative values (/ɪ/ longer than /i:/)	2	5%

Interestingly, by T2 (Figure 55) the results revealed the great majority of speakers obtaining positive outcomes. In fact, almost all of the speakers (except for two) produced the /i:/ longer than the /ɪ/ with an average length for the production of /i:/ of 332 ms and for /ɪ/ 243 ms. On the contrary, two speakers who produced a reversed length distinction did it with an average for /i:/ of 272 ms and /ɪ/ with an average of 281 ms.

**Figure 56**

*Duration difference between /i:/ and /ɪ/ at Time 3*



**Table 40**

*Summary T3*

<i>Summary T3</i>	N	%
a) Positive values	8	20%
b) Negative values (/ɪ/ longer than /i:/)	32	80%

However, striking results were found by T3 (Figure 56). At this time, the duration of /ɪ/ increased, and the duration of /i:/ decreased. That is, instead of maintaining the duration difference produced by T2 the results reveal the opposite tendency. 32 speakers produced the /ɪ/ with a mean duration for /ɪ/ of 243 ms and the /i:/ of 215 ms. Eight participants produced positive values with a mean for /i:/ of 231 ms and for /ɪ/ of 206 ms. These results suggest that for the front pair of vowels, most of the subjects did not manage to produce the correct quantity distinction over a year's time.

Overall, the results show that 80% of the speakers did not make a vowel quantity distinction. The final results (/i:/215 ms and /ɪ/ 243 ms) when compared with the RP references (/i:/ 293 ms and /ɪ/ 139 ms) show that the speakers did not shift towards the English norms.

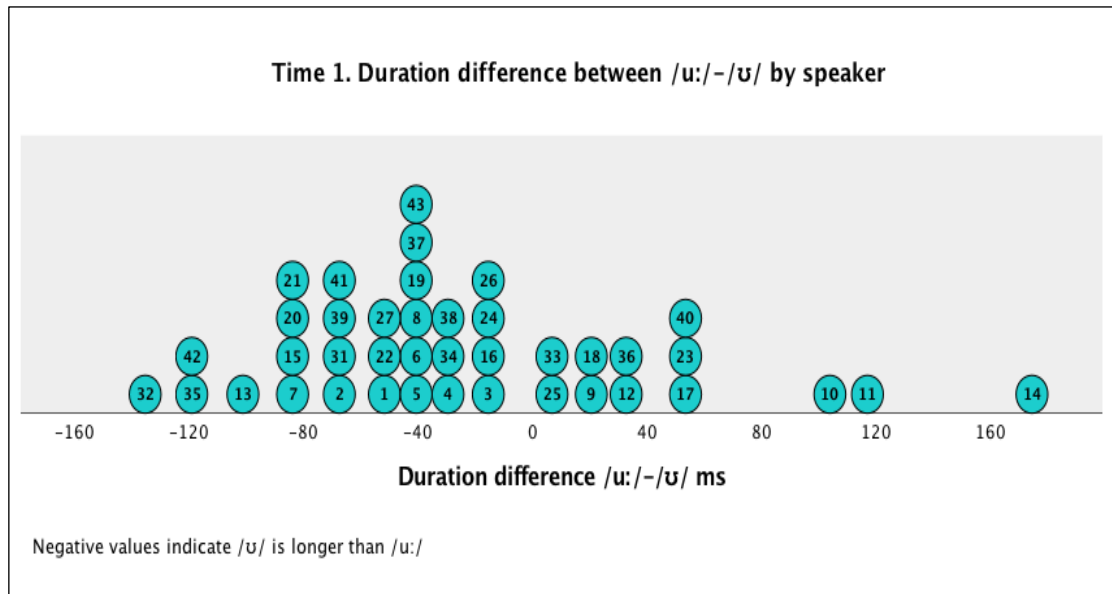


### 8.6.2 /u:/ and /ʊ/ vowels

This section shows the results for the length distinction (ms) between /u:/ and /ʊ/. Only vowels followed by voiced consonants were analysed.

**Figure 57**

*Duration difference between /u:/ and /ʊ/ at Time 1*



**Table 41**

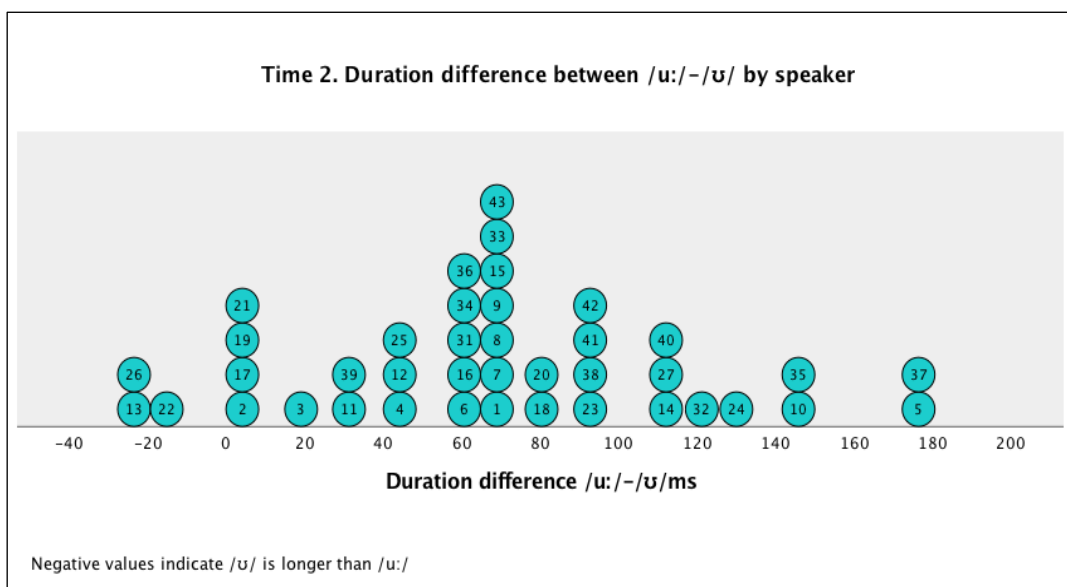
*Summary T1*

<i>Summary T1</i>	N	%
a) Positive values	12	30%
b) Negative values (/ʊ/ longer than /u:/)	28	70%

Figure 57 provides a clear illustration of speakers' distribution at the moment of producing the length contrast between /u:/ and /ʊ/. Results show that 28 participants produced /ʊ/ longer than /u:/ with an average duration for /u:/ of 210 ms and /ʊ/ 273 ms. Twelve participants managed the opposite length pattern with an average duration for /u:/ of 283 ms and for /ʊ/ of 224 ms.

**Figure 58**

*Duration difference between /u:/ and /ʊ/ at Time 2*



**Table 42**

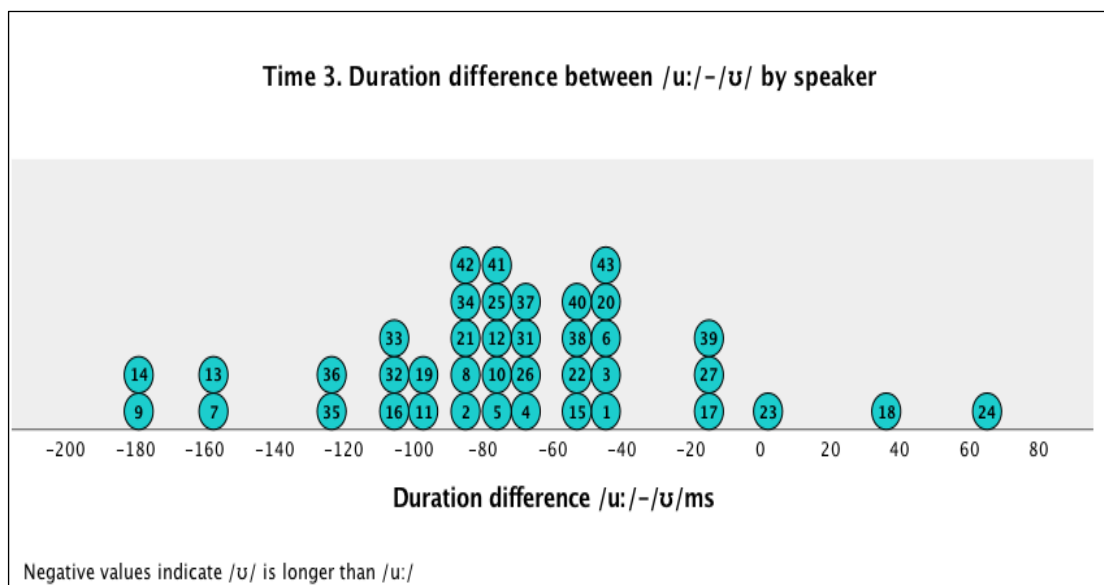
*Summary T2*

<i>Summary T2</i>	<b>N</b>	<b>%</b>
a) Positive values	37	92.5%
b) Negative values (/ʊ/ longer than /u:/)	3	7.5%

As opposed to T1, by T2 most speakers (37), as seen in Table 42, managed to produce the correct contrast duration. The results show an increase in the length of /u:/ with an average of 302 ms and a decrease in the duration of /ʊ/ with an average of 224 ms. Only three participants kept the negative values produced at T1 with an average for /u:/ of 252 ms and for /ʊ/ of 281 ms.

**Figure 59**

*Duration difference between /u:/ and /ʊ/ at Time 3*



**Table 43**

*Summary T3*

<i>Summary T3</i>	<b>N</b>	<b>%</b>
a) Positive values	3	7.5%
b) Negative values (/ʊ/ longer than /u:/)	37	92.5%

Results by T3 (Figure 59) show that 37 participants who previously (by T2) produced the correct length distinction between the vowels by T3 had reverted their values to their earlier realisations: that is, they went back to producing /ʊ/ with a mean of 323 ms and /u:/ with a mean of 242 ms. The three participants who were the exception to this trend produced vowels with a mean duration for /u:/ of 313 ms and /ʊ/ of 271 ms.

Overall, over a year, none of the speakers were unable to produce a quantity difference between the vowel pair. These results indicate that participants did not shift towards the British English norm, which was set for /u:/ at 294 ms and /ʊ/ with 142 ms.

In summary, by comparing the results between /i:/ and /ɪ/ and /u:/ and /ʊ/, it can be seen that most of the speakers did not produce a native/like quantity distinction between the pair of vowels examined. Interestingly, they followed the same pattern for the two pairs. That is, by

T1, 21 speakers who produced a longer /ɪ/ also produced a longer /ʊ/. By T2, 32 subjects who favoured a long /i:/ also did it for /u:/. And by T3, the 30 participants that incorrectly increased the length for /ɪ/ also did it for /ʊ/. These outcomes suggest a clear pattern in which the subjects were unable to produce quantity differentiation between English 'short' and 'long' vowels.

This section has shown the results for the quantity distinction between vowels. The subsequent section moves on to consider the results by providing the final chapter summary.

## 8.7 Summary

The aim of this chapter was to assess the progress across speakers in developing contrasts in the English vowel pairs /i:/ and /ɪ/, /u:/ and /ʊ/ and /e/ with respect to both spectral (vowel quality) and temporal (duration) features in running speech.

The results were compared to the contemporary RP reference values specified in Chapter 5 to observe whether the participants (individually) were shifting or adapting to the British English norms.

### 8.7.1 Formant analysis (vowel quality)

To summarise individual performance at the beginning and end of the year as measured by F1 - F2 Euclidean distances:

*/i:/ vs /ɪ/*

- Only 1 speaker began the year with a near- to- native -RP norm distinction between members of this pair (over 250 Hz). He/she still had a clear distinction at the end of the year.
- Only one speaker who did not have a near-native distinction at the start of the year had progressed to making one by the end.
- The other 38 speakers, who did not have a near-native norm distinction at the beginning of the year, had not developed one by the end.

### **/ɪ/ vs /e/**

- Sixteen speakers had a near- L1 distinction between this pair at the start of the year (over 250 Hz).
- By the end of the year these 16 speakers had maintained this distinction.
- Also, by the end of the year, they were joined by a further 23 speakers, who did not have a near- native distinction at the start of the year but had since developed one.

### **/u:/ vs /ʊ/**

- Thirty-two speakers had a near native distinction between this pair at the start of the year (over 110 Hz).
- By the end of the year 31 speakers who initially had a near-native distinction maintained it until the end of the year.
- None of the other 7 speakers, who did not have a near-RP norm distinction at the beginning of the year, had not developed one by the end.

### **/u:/-fronting**

To summarise individual performance between the beginning and end of the year as measured by F2 values:

- The average realisation for the group at T1 was 1547 Hz.
- Thirty-five speakers had already developed a fronted realisation by the start of the year (average F2 for this group = 1589 Hz).
- The remaining 5 speakers had not (average F2 for this group = 1253 Hz).
- By the end of the year 38 speakers produced non-fronted variants (average F2 for this group, which includes those who did not front initially and those who became non-fronters = 1262 Hz, i.e., F2 had been reduced by 20 Hz or more).
- By the end of the year only two speakers maintained their initial fronted variants (average F2 = 1519 Hz).

## **/ʊ/-fronting**

To summarise individual performance at the beginning and end of the year as measured by F2 values:

- The average realisation for the group at T1 was 1393 Hz.
- Twenty-seven speakers had already developed a fronted realisation by the start of the year (average F2 for this group = 1470 Hz).
- The remaining 13 speakers had not (average F2 for this group = 1231 Hz).
- Of the 27 speakers who fronted at the beginning of the year, 14 did less so by the end (i.e., F2 had been reduced by 20 Hz or more).
- Nine speakers who began the year as non-fronters were producing more fronted realisations by the end (average F2 at T3 for this group = 1376 Hz).

In respect of the beginning to end of year patterns for the vowel quality data, it is of interest that, as seen in the preceding sections, for some vowels and pairs, the development is not unidirectional over the year's course. For example, the T2 data show that some participants had not progressed between T1 and T2, and some had actually regressed before progressing again towards the native English values by T3.

### **8.7.2 Vowel duration**

- At the start of the year, the participants did not make a clear quantity distinction between /i:/ and /ɪ/, and /u:/ and /ʊ/. Some speakers, in fact, produced the short members of these pairs fractionally longer than the long members; others made the distinction in the direction associated with native English, but the durational differences in all cases only involved only 9 - 59 ms.
- By the end of the year, the situation remained much the same.

### 8.7.3 General

This chapter has provided the different degrees of separation that each individual speaker obtained across time points between contrasts in the English vowel pairs /i:/ and /ɪ/, /ɪ/ and /e/ and /u:/ and /ʊ/ with respect to both spectral and temporal features in running speech.

A discussion of the findings set out in this chapter is deferred until the final Discussion chapter.

The next Chapter examines the possible reasons for the differences between high- and low-performing individuals. In so doing, information from the language background questionnaire is drawn upon.

## **Chapter 9. Linguistic and social experiences in L2**

### **Language background questionnaire**

#### **9.1 Introduction**

A language background questionnaire was designed and administered to participants in order to determine which kinds of L2 experience and spoken interaction were most associated with greatest progress in developing contrasts between members of the 3 vowel pairs.

According to Flege (2018), the quality and quantity of L2 input is an essential aspect of successful L2 speech learning; what is more, speakers who have opportunities and the need to use English (L2) consistently will benefit from the development of their English speech (Flege & Liu, 2001). In this study, to track the quality and quantity of L2 input, data related to self-estimated use of English were collected through a language background questionnaire (see Appendix D) to observe possible changes in participants' social and/or academic behaviour and experiences at three timepoints in a year. As described in Chapter 4, the questions of the written questionnaire were divided into four categories:

1) background information, 2) attitudes towards English, 3) opportunities to develop the English language, and 4) circumstances of exposure to English. It should be borne in mind that the participants are all adults (over 18 years old) beyond the sensitive or critical period. Therefore, the overall results shed interesting light on the different patterns that speakers followed in their language use which correlate to the linguistic trends observed in changes in the production data.

##### **9.1.1 Structure of the chapter**

The sections of the chapter are as follows:

- summary of patterns
- results
- final summary



## 9.2 Summary of the patterns shown by the participants

The results of the acoustic data presented in Chapter 6 (citation-style) and Chapter 8 (read passage speech) have shown patterns about the number of speakers who had progressed (over the course of a year) to make an RP norm-like distinction between /i:/ and /ɪ/, /u:/and /ʊ/ and /ɪ/ and /e/ by means of spectral features. Contrasting duration differences were not developed by the speakers.

The progress patterns are as follows:

### 9.2.1 Word List Data – High Group

**Table 44**

*Word list*

	<b>All pairs</b>	<b>/i:/vs /ɪ/</b>	<b>/ɪ/ vs /e/</b>	<b>/u:/ vs /ʊ/</b>
<b>High performers</b>	1	5	6	25
<b>Low performers</b>	0	4	19	4

In citation-style, only one participant progressed to develop a native-like distinction between all the vowel pairs.

In citation-style, the same three speakers who developed a native-like distinction between /u:/and /ʊ/ also did it for /ɪ/ and /e/.

In citation-style, the same four speakers who developed a native-like distinction between /i:/and /ɪ/ also did it for /u:/and /ʊ/.

In citation-style, one participant progressed to develop a native-like distinction between /i:/ and /ɪ/ and /ɪ/ and /e/.

## 9.2.2 Connected Speech Data - High Group

**Table 45**

*Connected Speech*

	<b>All pairs</b>	<b>/i:/vs /ɪ/</b>	<b>/ɪ/ vs /e/</b>	<b>/u:/ vs /ʊ/</b>
<b>High performers</b>	0	1	3	4
<b>Low performers</b>	1	1	36	27

None of the same speakers (in the word list data and scripted passages) progressed to develop a native-like distinction between all the vowel pairs examined.

None of the same speakers produced a native-like distinction between all vowel pairs in reading scripted passages.

Based on the above patterns, six participants were identified to be the same ones who developed a clear distinction between /i:/ and /ɪ/, /u:/ and /ʊ/, /ɪ/ and /e/ (above the RP norm threshold). In addition, eight speakers were identified to be the same individuals who did not make progress between these vowels over a year. These speakers were compared against each other by means of their language background questionnaire responses to determine which types of exposure and which contexts were most conducive to their different performances.

## 9.3 Results

The results of the language background questionnaire were divided into a high performing group and a low performing group.

It is important to note that, at T1, the six high performing speakers did not present the targeted separation between the vowel pairs. They developed the contrast of vowels from T2 onwards.

### 9.3.1 High performing group

The results are presented following the sections of the language background questionnaire:

**a) Attitudes toward English:** in this section subjects were asked to report their confidence level at the moment of performing the four English language skills (listening, speaking, reading, and writing), as well as declaring any difficulties in understanding British English when interacting with native speakers. Also, this part of the questionnaire asked about the motivation to speak in English, the importance of achieving native-like pronunciation, the importance of maintaining their Spanish accent while speaking in English, and finally, the significance of developing the four English skills and of modelling the British accent.

The results from the high performing group show a tendency to report reading as their strongest skill throughout the year. By T2, there was a slight change regarding the confidence for writing, but overall, listening and reading (both receptive skills) obtained the highest scores. Speaking was not considered as a strong skill.

Moreover, when the participants were asked about the most difficult skill to perform, changes were also logged by T2, where 50% of the speakers reported listening and speaking as the most difficult skills to perform. Speaking was reported by 50% of the speakers as a difficult task to perform at the three times studied. Table 46 is a sample of the responses.

**Table 46**

*Self-reported most difficult skill*

Which of the 4 skills is the most difficult?	Listening	Speaking	Reading	Writing
Time 1	25%	50%	-	25%
<b>Time 2</b>	50%	50%	-	
Time 3	25%	50%		25%
<b>Change: T1-T2</b>	<b>(+)25%</b>			<b>(-)25%</b>
<b>Change: T2-T3</b>	<b>(-)25%</b>			<b>(+)25%</b>

Regarding the level of difficulty of understanding British English, the results show a progress in the understanding of them during the year, ranging from: 'very difficult' (50%) at T1, to 'slightly difficult' (50%) by T2, and 'not difficult' (50%) by T3.

These outcomes suggest an improvement in the language comprehension after being heavily exposed to British English.

In terms of motivation, the speakers did not present changes during the times examined. All of them reported being highly motivated throughout the year. Attitudes towards achieving native-like pronunciation and maintaining the Spanish accent while speaking in English did undergo some changes by T2 and T3. Responses related to attaining a native-like English accent decreased from 'extremely important' (75%) at T1, to 'very important' by T2 and T3 (50% respectively). Although there was a change in the responses, the tendency was still within the 'important' value.

Moreover, for the speakers of this group, maintaining their Spanish accent while speaking in English was not important, and in terms of the variety of English they were hoping to achieve at T1, 50% of the speakers stated, 'American English' as their target accent; however, during T2 their responses changed to 'British English' (100%).

**b) Opportunities to develop the English language:** this section of the questionnaire asked respondents to give information related to their use of English in academic<sup>17</sup> and non-academic contexts. When asked about the time spent at the university in different academic settings such as seminar classes, supervision meetings, and tutorial time, the results reveal interesting variations between the times examined. Table 47 is a sample of the questions and responses.

---

<sup>17</sup> Changes in attendance in academic settings may or may not have been personal choice but determined by changes in course requirements over the year.

**Table 47***Time spent at the university in academic activities*

At the University of York, how many hours do you spend at the university?				
Hrs:	7 to 9	4 to 6	1 to 3	0
Time 1	50%	-	50%	-
<b>Time 2</b>	75%	-	25%	-
Time 3	25%	25%		25%
<b>Change: T1-T2</b>	<b>(+)25%</b>		<b>(-)25%</b>	
<b>Change: T2-T3</b>	<b>(-)25%</b>	<b>(+)25%</b>		<b>(+)25%</b>

The above results show that by T2, participants increased their hours and time spent at the university interacting in academic settings, although some hours of interactions were reduced by T3.

In response to the question related to the interaction with native British English speakers, the outcomes present a favourable progress tendency during the times observed. At T1, the ‘not often’ (50%) result changed to ‘sometimes’ (75%) by T2, and to ‘very often’ (25%) and ‘sometimes’ (75%) by T3. These results show a positive increment in the interactions between the participants and native English speakers. Furthermore, questions concerning participation in extracurricular activities involving native English speakers, and work off/on campus, did not obtain positive results. Participants did not work on/off campus and they did not participate in extracurricular activities.

**c) Circumstances of exposure to English:** questions related to living arrangements did not present changes during the year. None of the speakers declared to live either with British host-families, or with their Spanish-speaking family in the UK. Moreover, by T2, 50% of the participants declared to live with their Spanish-speaking partner, a situation that changed (decreased) to 25% by T3.

When asked about house-sharing with native British English speakers, some differences were observed. At T1, 50% of the respondents declared sharing a house with British native speakers. That specific percentage increased by T2 reaching 75%, and later at T3, decreased

to 25%. Moreover, concerning close friendships with native British speakers 75% of the speakers stated a negative answer.

In the final part of the survey and only during T3, participants were asked to determine and self-evaluate their English language use over a year. Three out of four speakers declared to speak more Spanish than English (75% vs 25%) during a day; and in terms of speaker interaction, 70% of the participants indicated that they spoke English mostly with native British English speakers, whereas the other 30% of the speakers reported to use their English language with international students (not native English speakers).

Taken together, these results suggest a possible association between the participants, the exposure to English spoken by native speakers, and the English language used, performed and experienced by the speakers during the course of a year.

The six participants that managed to make progress (from T2 onwards) in their vowel contrast separation between /i:/and /ɪ/, /u:/ and /ʊ/ and /ɪ/ and /e/ also reported progress and improvement in their attitudes towards the English language, the opportunities to speak English with native speakers (academic and non-academic settings), and improvement in the exposure to English spoken by native speakers from T2 and beyond. Thus, changes in the quality and quantity of the English input could well have had an impact on the results of the vowel separation produced by the participants, since both changes (vowel distinction and exposure/interactions) occurred from T2 onwards.

### 9.3.2 Low performing group

The low performing group was made up of eight speakers. These participants did not produce a clear distinction between any of the pairs of vowels (/i:/-/ɪ/, /ɪ/-/e/, /u:/-/ʊ/) above the RP norm threshold during the three times observed. The results of the language background questionnaire are the following:

**a) Attitudes toward English:** the results of participants' confidence level at the moment of performing the four English language skills (listening, speaking, reading, and writing) presented a trend towards reporting the receptive reading skill as the strongest one throughout the year, although by T2 and T3 slight changes were recorded regarding the confidence level towards writing and listening skills, which each obtained a 12% improvement when compared to T1.

Interestingly, when the subjects were asked about the most difficult skill to perform, the speaking (productive) skill response increased over the times tested, going from 15% at T1, to 38% by T2 and 50% by T3. In addition, the listening skill was reported as the second most difficult one to achieve. The other two abilities did not undergo important changes during the year.

**Table 48**

*Self-reported most difficult skill*

Which of the 4 skills is the most difficult?	Listening	Speaking	Reading	Writing
Time 1	25%	15%	13%	38%
Time 2	38%	38%	12%	12%
Time 3	25%	50%	12%	12%
<b>Change: T1-T2</b>	<b>(+)13%</b>	<b>(+)23%</b>		<b>(-)26%</b>
<b>Change: T2-T3</b>	<b>(-)13%</b>	<b>(+)12%</b>		

In terms of the level of difficulty in understanding British English varieties, the outcomes show changes from T1 (50% not difficult) to T3 (50% slightly difficult), showing less improvement or greater realism in language comprehension towards the end of the year. Moreover, regarding motivation to speak English, the speakers did not present variations during the times examined. 87% of the participants were highly motivated and 12% were very motivated, both outcomes showing positive attitudes to the English language. Likewise, favourable attitudes about achieving native-like pronunciation were self-reported by the group showing a trend leaning towards ‘extremely important’ (75%) and ‘very important’ (12%). Surprisingly, when reporting how important it was to them to maintain a Spanish accent while speaking English, the group had responses of ‘not important’ (75%) and ‘very important’ (25%), the latter being at odds with their statements for the previous question (native-like pronunciation). Finally, in terms of the variety of English they wanted to achieve, results show that speakers mainly targeted British English (87%) rather than American English (13%).

**b) Opportunities to develop the English language:** the responses to questions related to the use of English in academic and non-academic contexts such as seminar classes, supervision meetings, and tutoring time show some variations, especially by T2.

**Table 49**

*Time spent at the university in academic activities*

At the University of York, how many hours do you spend at the university?				
Hrs	7 to 9	4 to 6	1 to 3	0
Time 1	87%	13%	-	-
Time 2	63%	25%	12%	-
Time 3	63%	25%	12%	-
<b>Change: T1-T2</b>	<b>(+)24%</b>	<b>(+)12%</b>	<b>(+)12%</b>	
<b>Change: T2-T3</b>				

By T2, students increased their exposure to English in academic contexts. They reported more hours of tutorial time, supervision meetings, and activities within the university. This heavily exposed period was reduced at T3, reflecting the percentages specified at the beginning of the year (T1).

In terms of the interaction with native British English speakers, responses show favourable progress at T1, with 50% of the speakers reporting ‘sometimes’ results that increased at T2 to ‘very often’ 63%. However, important changes were observed at T3. The ‘very often’ interaction dropped to 37%, the ‘sometimes’ option decreased to 38%, and the ‘not often’ response increased to 25%. These results show a possible lack of interaction with native speakers in non-academic settings.

Additionally, as for taking part in extracurricular activities involving native English speakers, mixed responses were obtained. 50% of the speakers reported participating in non-academic activities, while the other 50% reported they did not. Regarding work conducted off/on campus, the results show that 75% of the speakers did not work whereas 25% did, and the multiple accents to which the participants (involved in non-academic activities and work) were exposed during their interactions was British English.



**c) Circumstances of exposure to English:** questions related to living arrangements did not exhibit changes during the year. None of the speakers declared living either with British host-families or with their Spanish speaking family in the UK, and only 13% declared living with his/her Spanish-speaking partner. However, when asked about house-sharing with native British English speakers some differences were recorded. At T1, 75% of the respondents declared sharing a house with British native speakers. That percentage decreased by T2 to 62% and by T3 to 50%. Moreover, concerning close friendships with native British speakers, responses varied over the course of a year. At T1, 85% of the speakers stated having close native English-speakers friends, while by T2 the percentage had decreased to 50%, and by T3 the responses had increased to 75%. However, variations were also observed among speakers who did not report close friendships with native English speakers. The results show that at T1, 13% of the participants declared not having close native English-speakers friends, while at T2 the percentage increased to 50% and that by T3 it had decreased to 25%.

Lastly, at the end of the questionnaire and only during T3, participants were asked to determine and self-assess their use of the English language during the year. Interestingly, the results show that 87% of the participants considered speaking more English than Spanish during the day; however, 63% of the speakers reported that these English interactions were with non-native English speakers while 37% of the participants reported that they were with native English speakers.

Overall, these results show that in terms of attitudes toward English, the participants declared the receptive reading skill as the least difficult to perform, whereas speaking was considered the most difficult skill; the latter result increased in terms of percentages of participants who stated that speaking became more challenging over time.

In addition, listening comprehension did not improve over the times examined. These outcomes could be reflected in the lack of progress in producing distinctive vowels.

Moreover, their exposure to academic and non-academic English increased mainly during T2, and apparently, close friendships and circumstances to have direct contact with native English speakers were registered primarily at T1. In terms of daily English use, participants reported speaking English most of the time in a day. However, these exchanges occurred principally with international students whose first language was not English.

## 9.4 Summary

In summary, comparing the two sets of results (high performing vs low performing group), it can be said that the participants experienced some similarities but also important differences in the quality and quantity of English input received for one year. These variations seem to be more predominant in the aspects of ‘opportunities to develop English’ and ‘circumstances of exposure to English’.

The following Table presents a summary of the main results.

**Table 50***Summary table*

<b>Similarities ✓</b>	<b>Differences X</b>
<b>a) Attitudes toward English</b>	<b>a) Attitudes toward English</b>
<ul style="list-style-type: none"> <li>✓ Highly motivated to speak English</li> <li>✓ To achieve native-like British pronunciation was a goal</li> <li>✓ To maintain their Spanish accent while speaking was not important</li> <li>✓ Speaking was the most difficult English skill to perform</li> </ul>	<ul style="list-style-type: none"> <li>X The high performers reported progress in understanding the different British English accents. The low performers did not.</li> <li>X The high performers found the speaking skill as the most difficult. The low performers did not.</li> </ul>
<b>b) Opportunities to develop the English language</b>	<b>b) Opportunities to develop the English language</b>
<ul style="list-style-type: none"> <li>✓ Both groups had more exposure to English in academic and non-academic settings by Time 2.</li> </ul>	<ul style="list-style-type: none"> <li>X By T3, the high performers increased their interactions in academic and non-academic settings with native speakers. The low performers reduced them.</li> <li>X The high performers did not participate in extracurricular activities. The low performers did.</li> </ul>
<b>c) Circumstances of exposure to English</b>	<b>c) Circumstances of exposure to English</b>
<ul style="list-style-type: none"> <li>✓ Neither group lived with British host-families nor with their Spanish speaking family in the UK</li> <li>✓ None of the members of the groups have native English speakers as partners</li> <li>✓ None of the members of the groups have native English-speakers as close friends</li> </ul>	<ul style="list-style-type: none"> <li>X The high performers spoke more Spanish than English during the day. The low performers did the opposite.</li> <li>X The high performers reported speaking English with native speakers most of the time. The low performers spoke English mostly with international speakers (e.g., Chinese, Italian, or other Spanish speakers), showing a difference in the quality of input rather than quantity.</li> </ul>

The above comparative table provides the reader with the main differences and similarities between both groups at a glance. As noted, different linguistic and social experiences were identified as influencing the development of the English vowel contrast.

The social factors which appeared to exert a positive effect on the participants who managed to progress in the production of a native-like contrast between the pair of vowels were maintaining and increasing academic and non-academic social interactions with native English speakers after T2. These results agree with the SLM, L2LP and PAM models (discussed in Chapter 3) that emphasises the importance of the quality and quantity of the L2 input. The fact that the high performing group had most of their interactions with native British English speakers influenced their understanding of English, which certainly shaped and led to the progress of their vowel productions. By contrast, the pattern shown by the low performing group indicates that the possible 'low quality' of English input (non-native speakers) and the decrease of academic activities after T2 played a central role in the lack of English understanding and, therefore, in the production of English vowel contrasts.

Contrary to expectations, some differences between high and low performers' responses were found to be related to the confidence at speaking in English, where unexpectedly, the high performers reported speaking as a difficult skill to perform throughout the year, whereas low performers did not. Others surprising results were in terms of the L1 use during a normal day, where unexpectedly, the high performing participants reported to speak more Spanish than the low performing ones over the course of a day.

I return to the interpretation of these results in the Discussion chapter (Chapter 10). .

## Chapter 10. Discussion

### 10.1 Introduction

The purpose of this research was to examine the progress of adult Spanish speakers, both individually and as a group, in distinguishing between English vowels /i:/ and /ɪ/, /ɪ/ and /e/ and /u:/ and /ʊ/ in terms of both spectral and duration features; and to identify and consider the external factors associated with the different rate of progress among speakers.

This chapter presents a general discussion of the overall findings by evaluating the results across the chapters against the six research questions and hypotheses set out in Chapter 1.

#### 10.1.2 Structure of the chapter

The sections of the chapter are as follows:

- discussion of the findings
- limitations of the research and future directions

### 10.2 Discussion of the findings

For convenience of reference, the research questions (bold type) and associated hypotheses (italics) are repeated (partially in paraphrase) below:

**Q1. Does the group as a whole show a movement towards producing phonemic contrast between /i:/ and /ɪ/, /ɪ/ and /e/ and /u:/ and /ʊ/ over the course of their year in England?**

*(H1) It was hypothesised that, given the immersion of the group in a range of English social, commercial and educational linguistic settings, one would expect to find adaptation to L1 patterns, as manifest in these vowel contrasts and pronunciations.*

#### Findings:

The findings for /i:/ and /ɪ/ and for /u:/ and /ʊ/ show that the group as a whole did show some movement towards producing a phonemic contrast between these two pairs of vowels; however, by the end of the year the distinctions made by the group fell short of native RP English norms. The findings for /e/ and the /ɪ/ - /e/ contrast show that the group generally moved towards a native English-like pronunciation of /e/ and made a clear native-like distinction between the /ɪ/ - /e/ pair by the end of the year.

The phonemic contrast /i:/ and /ɪ/ at T1 was realised with an average of 112 Hz. This value by T3 increased to 235 Hz, but it was still a little below the English norm.

These findings further support the previous position of some scholars (e.g., Flege et al., 1992 & Guion et al. 2001) that states that L2 immersion by means of residency, especially for adult learners, helps learners to improve the performance of their L2. In addition, the findings for /i:/ and /ɪ/ are in line with previous studies by, for example, Flege et al. (1997), Fullana-Rivera & Mackay (2003), and Morrison (2002), which have shown a lack of distinction between /i:/ and /ɪ/ by Spanish L2 learners. Their results are mainly associated with a short time of exposure (less than a year) to the new vowel categories, which differ from the group's immersion length with an average separation of 268 Hz, which increased by T3 to 394 Hz.

**Q2. If the Spanish learners do mark the contrasts between the three vowel pairs, do they do so by vowel quality, as indexed by the first and second formant values, or by length, or both?**

*(H2) It was anticipated that, given the findings of certain previous studies of Spanish learners, which have found movement towards native English vowel quality over shorter periods of time (see Munro & Derwing, 2008 and Flege, 1988), one would expect to find - minimally - changes in quality; there was no hypothesis as to whether the present participants would also begin to mark the contrasts durationally.*

#### Findings:

*Duration:* differences in vowel length between members of the front and the back vowel pairs did not occur by end of the year. As a group, participants produced the length for /i:/ and /ɪ/ the same (on average 438 ms), and /u:/ and /ʊ/ with very similar durations (/u:/ - 449 ms, and /ʊ/ 441 ms). This result supports the position that speakers of languages that do not use durational distinctions to show vowel contrast would have problems in adapting their productions on the temporal dimension (Escudero, 2000; Kim et al., 2018; Kondaurova & Francis, 2010). With hindsight, this could perhaps have been hypothesised at the outset.

*Quality:* The phonemic contrast /i:/ and /ɪ/ at T1 was realised with an average of 112 Hz; this value by T3 increased to 235 Hz. It was below the English norm for the quality distinction, but not by a very long mark. Moreover, if the RP English norm is 250 Hz and the T3 norm for this group is 235 Hz, there is likely to be some overlap between the two populations. For the

/i:/ - /ɪ/ contrast, the results showed rather more movement in the vowel quality of the vowel pair. The group produced /i:/ and /ɪ/ with higher F1 and F2 values than the English norm, obtaining a lower and more fronted position than the English targets. The final results for /ɪ/ showed an increase in the F1 value (407 Hz) which shows a movement of /ɪ/ to a more open position over the course of a year.

The fact that the development of the vowel pair contrast was marked by the production of the /ɪ/ in a more open position signalled a progress in the use of spectral cues to produce English vowels. According to the SLM, speakers could have created a new sound (after a year in England) because contrary to the assimilation pattern presented by Flege (1991), and the results of Fullana-Rivera & Mackay (2003), the production of /ɪ/ was not assimilated into the Spanish /i/ category, which has a F1 value of 286 Hz approx.

The developing contrast between /u:/ and /ʊ/ was marked by vowel quality, as indexed for /u:/ in F2 and for /ʊ/ in F1 and F2. The final result for /u:/ shows a statistical increase in the F2 value, resulting in a more fronted vowel than the one produced at the beginning of the testing period. The final result for /ʊ/ shows an increase in the F1 and F2 formants. The higher F1 value marked a movement of the /ʊ/ to a more open position, and the higher F2 value showed a progress towards a more fronted one.

These results suggest that the speakers created a contrast nearer to contemporary native English norms. The fronting of /u:/ and /ʊ/ show an accommodatory gravitation towards present-day /u:/ and /ʊ/ targets, shifting from very high back (English learning textbook-like) vowels to more central ones, possibly owing to a systemic change but also as a result of exposure to the newer English fronting tendency. The fact that the mean age of the group was 27 years old could have been a factor in the realisation of these tokens nearer to native English speakers, since /u:/ and /ʊ/ fronting is a phenomenon that younger speakers tend to exhibit more than older ones (Harrington et al., 2008; Hawkins & Midgley, 2005), and as a consequence of immersion and interaction with their native English-speaking peers, these ongoing changes were reflected in the group performance at the end of the year.

These findings are somehow contrary to previous studies, e.g., Escudero & Chládková (2010) because the changes in F1 and F2 values, especially the movement of the /ʊ/ to a more open and fronted position, demonstrated that the speakers created a new vowel category by dividing the Spanish /u/ into two, since /ʊ/ was produced with an average of 451 Hz (not similar to the Spanish value for /u/ of 322 Hz). Also, these findings differ from those of Koffi

& Lesniak (2019) and Wang & Munro (1999), which have suggested more than a year (slow process) for back vowel changes, because although the distinction between the pair was not greater than the English norm, the formants did change in a year period, so the process of adaptation was faster than the ones discussed by these previous authors.

The developing contrast between /ɪ/ and /e/ was marked by quality changes for both vowels. By T3, F1 values for /e/ decreased to around 600 Hz, and the F2 values increased above 1900 Hz. The F1 decrease indicates a more open vowel, possibly owing to a systemic ‘push’ effect, but also as an adjustment towards native speaker-norms (a vowel closer to Cardinal 3). This result does not support the findings and assimilation patterns found by Escudero & Chládková (2010) and Flege (1991), who have suggested that Spanish speakers’ realisation of the English /e/ will be assimilated to Spanish /e/, for which F1 values have an average of 458 Hz and F2 values have an average of 1814 Hz.

**Q3. Over time, do some individual speakers develop more marked contrasts than others, and, if so, what are the specific experiential factors associated with them?**

*(H3) It was hypothesised that, given the learners would have different degrees of exposure to, and experience of, English speaking contexts - educational, commercial and social - it was anticipated that there would be differential rates of progression. While it is possible to hypothesise at a general level that the amount of exposure to native English, and active engagement with native English speakers rather than simply passive exposure to the language, might be facilitative of progression, no hypothesis was put forward in respect of which specific contexts would be most associated with advanced progression.*

**Findings:**

Overall, the results show variations among individuals in the development of marked contrasts between the vowel pairs. These differences were produced by T2 onwards (after five months of living in England) which indicate a possible new timeframe for the L2 phonetic/phonological learning process, since the results of progress differ from previous claims suggesting that non-native English speakers in an immersion context will progress fast initially (0-5 months) and then plateau (see Flege et al.,1992). The individual results show that some learners develop marked vowel contrasts faster than others. Progress made by the same individuals with respect to all vowel pairs examined were only seen in the citation-style data,



which supports the use of this elicitation method to show something nearer to the underlying competence of the speakers. A total of six speakers developed a contrast (T2-T3) between all the pairs, and eight speakers did not produce a clear distinction between the pair during the whole year. The specific factors associated with this discrepancy in the speed of progress to produce vowel contrasts could be attributed to the ability to understand the English spoken by native speakers, the social interactions in academic and non-academic settings with native English speakers, and the use of the English language with native speakers versus international English speakers. These are all factors to which 'high performers' were much more engaged than the 'low performers'. These results support the claims by Flege (2018) and Jun & Cowie (2004), which indicate that adult learners can progress towards making a clear distinction between contrasting vowels by having an active involvement in spoken interaction with native English speakers, [hence], good quality and quantity of input, rather than a passive exposure to the L2. The fact that the 'low performers' experienced fewer social exchanges with native English speakers and more English-speaking interactions with international users of English may have played a part in this, as it could well amount to low input quality.

#### **Q4. Are the contrasts - as determined by acoustic measurements - perceived by native English speakers?**

*(H4) It was hypothesised that the contrasts will be perceived, and based on studies by, for example, Flege et al. (1997) and Hillenbrand et al. (2000), those that are marked by vowel quality will be more perceptually salient for L1 English speakers than those marked just by vowel length differences.*

#### **Findings:**

The results show that the vowel contrasts determined by the acoustic measurements – F1 and F2 values - were perceived by native speakers. The fact that L1 English speakers correctly identified vowels produced by participants with a clear vowel quality distinction and misidentified vowels produced by speakers without it, support the claims by, for example, Francis et al. (2000) and Zahorian & Jagharghi (1993), which state that native English speakers rely on spectral cues over temporal ones to distinguish English vowel contrasts. The misidentification patterns found in this research are consistent with those found by Fullana-Rivera & Mackay (2003) to the effect that Spanish speakers' productions of /ɪ/ were misidentified as /i:/ more than 50% of the time, and with Escudero & Chládková, (2010),

Flege (1991) and Flege et al. (1997), who have proposed assimilation patterns for /ɪ/ to /i:/ and sometimes to /e/. However, the misidentification responses for /u:/ and /ʊ/ do not support the findings from the previous authors, which indicate that both vowels are assimilated and produced (by Spanish speakers) as the Spanish /u/. The results for the group with no clear quality distinction between the back pair showed that English native speakers identified /u:/ as /ʌ/, and /ʊ/ as /æ/, both more open vowels.

**Q5. Do the patterns found in the present data have a bearing on the conceptual models of L2 phonological/phonetic learning?**

*(H5). It was hypothesised that given the fact that the different models share similarities and differences regarding L2 phonological/phonetic development. It is hypothesised that the patterns found in the present data will be related to the conceptual models, the emphasis of which is placed on adult's L2 production over children's L2 system development, and in models where learning English vowel contrast is specially addressed, as in the Speech Learning Model by Flege (1995) and the Second Language Linguistic Perception Model of Escudero (2005) (respectively).*

Findings:

The conceptual models presented in this research do have a bearing on the patterns found in the present data. At different levels and in different respects, all of them (CPH; SLM; L2LP; PAM) present a theoretical basis to support the results. All of the models support the idea that L2 learning is optimised during childhood. The different results for the participants of this study (all adults) bear out the notion that, although learning and adaptation towards native-like production of vowels is possible, this process takes time for adults (over five months of deep and active immersion). The movements of the short vowels to more open and fronted positions, without being assimilated to the Spanish production of vowels, reflect the possible creation of new categories as proposed by the SLM.

The results obtained at different time points tend to support the different stages for learning a new language proposed by the L2LP model. The specific duration results (which showed a lack of distinction on vowel duration) might reflect the possibility that participants were unable to progress beyond the 'initial stage' of learning proposed by this model, showing that in terms of the duration feature, the subjects did not develop a new uncategorised dimension. However, in terms of vowel quality, the different progression rates across participants indicate

the potential development of category boundaries at a different level during the L2 learning - as stated in the L2LP model. The results related to the different degrees of exposure to, and experience of, English speaking contexts - educational, commercial and social - that learners reported experiencing during a year in England, have shown and emphasised the importance of the quality and quantity of the L2 input to progression towards a native English type of vowel production, as theorised in SLM, L2LP and PAM models.

### **10.3 Limitations and future directions**

The development of English vowel contrasts by L2 speakers has been shown to be influenced by multiple factors such as age of learning (DeKeyser, 2000), learning environment (Best & Tyler, 2007), length of immersion (Guion et al., 2000), and use of the L1 and L2 (Polka, 1991), among others. One limitation of this study was the failure to track whether participants left the UK during the time tested (e.g., returned home for more than a month). This could be a useful piece of information related to time and use of English. Another limitation was the absence of a perception test (for native speakers) that included data from Time 1 to compare with Time 3. This was not possible due to the timeframe of this research and the COVID-19 pandemic restrictions.

This longitudinal study has contributed useful insights into L2 phonemic development. It has shown how adult non-native English speakers progress towards vowel contrast adaptation while immersed on the target language. Future longitudinal research can be conducted in segmental analysis by addressing different vowel such as /ɔ:/, /ʌ/, /æ/, /ɑ/ and testing them for a longer period. Also, in terms of suprasegmental analysis, future research might focus on of the development of native-like intonation patterns or rhythmic features.

Finally, future research that I plan to conduct might exploit the unanalysed data collected in this research, such as the more naturalistic picture description recordings, and the unscripted self-description recordings.

### **10.4 Supplementary information**

The raw data (un-normalized formants and duration values) in citation and connected speech style are available at [osf.io/jdxht](https://osf.io/jdxht)

## References

- Abercrombie, D. (1964). *English phonetic texts*. London: Faber and Faber Ltd.
- Adank, P., Smits, R., Van Hout, R. (2004). A comparison of vowel normalization procedures for language variation research. *The Journal of the Acoustical Society of America*, 116 (5) 3099-3107.
- Ali, E. (2013). Pronunciation problems: Acoustic analysis of the English vowels produced by Sudanese learners of English. *International Journal of English and Literature*, 495-507. [https://doi: 10.5897/IJEL12.031](https://doi.org/10.5897/IJEL12.031)
- Amorós-Negre, C. (2016). The spread of Castilian/Spanish in Spain and the Americas: A relatively successful language standardisation experience. *Sociolinguistica*, 30 (1) 25-44. <https://doi.org/10.1515/soci-2016-0003>
- Amstutz, A. (2020). The history of Spanish language [Blog post]. Retrieved from <https://lingvist.com/blog/history-of-spanish/>. Last accessed January 2023
- Audacity. (2021). (Version 3.0.2) [Computer software]. Retrieved from <https://www.audacityteam.org>
- Baker, W., & Trofimovich, P. (2006). Perceptual paths to accurate production of L2 vowels: The role of individual differences. *International Review of Applied Linguistics in Language Teaching*, 44 (33) 231-250 [https://doi:10.1515/IRAL.2006.010](https://doi.org/10.1515/IRAL.2006.010)
- Balota, D., Yap, M., Cortese, M., Hutchison, K., Kessler, B., Loftis, B., Neely, J., Nelson, D., Simpson, G., & Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods*, 39, 445-459.
- Baptista, B. O. (2006). Adult phonetic learning of a second language vowel system. In B. O. Baptista (Ed.), *English with a Latin beat: Studies in Portuguese/ Spanish-English interlanguage* (pp.19–40). Amsterdam: Benjamins.
- Bauer, L. (1985). Tracing phonetic change in the received pronunciation of British English. *Journal of Phonetics*, 13, 61–81.
- Best, C. (1995). A direct realist view of cross-language speech. In W. Strange (Ed.), *Speech perception and linguistic experience* (pp. 171–204). Baltimore: York Press.
- Best, C., & Strange, W. (1992). Effects of phonological and phonetic factors on cross-language perception of approximants. *Journal of Phonetics*, 20, 305-330.
- Best, C., & Tyler, M. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In O.-S. Bohn & M. J. Munro (Eds.), *Language experience in second language speech learning* (pp. 13–34). Amsterdam: Benjamins.

- Bjelaković, A. (2017). The vowels of contemporary RP: Vowel formant measurements for BBC newsreaders. *English Language and Linguistics*, 21(3), 501-532.  
[https://doi:10.1017/S1360674316000253](https://doi.org/10.1017/S1360674316000253)
- Boersma P. (1998). *Functional Phonology*. [Doctoral dissertation, University of Amsterdam]. The Hague: Holland Academic Graphics.
- Boersma, P., & Weenink, D. (2023). Praat: doing phonetics by computer. (Version 6.0.49) [computer program]. Retrieved 2023 from <http://www.praat.org>
- Bohn, O-S. (1995). Cross-language speech perception in adults: First language transfer doesn't tell it all. In W. Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues in cross-language speech research* (pp. 279–304). Baltimore: York Press.
- Cano, R. (1995). La historia del español. In Manuel Seco & Gregorio Salvador (eds.), *La lengua española, hoy*, (pp.23–35). Madrid: Fundación Juan March.
- Casillas, J. (2020). Phonetic category formation is perceptually driven during the early stages of adult L2 development. *Language and Speech*, 63 (3) 550-558.
- Cebrian, J. (2006). Experience and the use of non-native duration in L2 vowel categorization. *Journal of Phonetics*, 34, 372–387.
- Chládková, K., Escudero, P., & Boersma, P. (2011). Context-specific acoustic differences between Peruvian and Iberian Spanish vowels. *The Journal of the Acoustical Society of America*, 130 (1).
- Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, Massachusetts: MIT Press.
- Clopper, C. (2009). Computational methods for normalizing acoustic vowel data for talker differences. *Language and Linguistics Compass*, 3 (6), 1430-1442.
- Colantoni, L., & Gurlekian, J. (2004). Convergence and intonation: historical evidence from Buenos Aires Spanish. *Bilingualism: Language and Cognition*, 7 (2) 107-199.  
[https://doi: 10.1017/S1366728904001488](https://doi.org/10.1017/S1366728904001488)
- Council of Europe. (2001). *The Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge: Cambridge University Press.
- Coupland, N., Coupland, J., Giles, H., & Henwood, K. (1988). Accommodating the elderly: Invoking and extending a theory. *Language in Society*, 17 (1), 1-41.  
[https://doi:10.1017/S0047404500012574](https://doi.org/10.1017/S0047404500012574)
- DeKeyser, R. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition*, 22, 499-533.

- Deterding, D. (1997). The formants of monophthong vowels in standard southern British English pronunciation. *Journal of the International Phonetic Association*, 27, 47–55.
- Drljača, B. (2017). Communication courtesy or condescension? Linguistic accommodation of native to non-native speakers of English. *Journal of English as a Lingua Franca*, 6 (1), 29-55. <https://doi.org/10.1515/jelf-2017-0006>
- Encyclopædia Britannica. (2021). Spanish language [image]. <https://www.britannica.com/topic/Spanish-language#/media/1/558113/239454>
- Escudero, P. (2000). *Developmental patterns in the adult L2 acquisition of new contrasts: The acoustic cue weighting in the perception of Scottish tense/lax vowels by Spanish speakers*. [Unpublished master's thesis, University of Edinburgh]. Edinburgh.
- 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. (2005). *Linguistic Perception and Second Language Acquisition: Explaining the Attainment of Optimal Phonological Categorization*. [Doctoral dissertation, Utrecht University]. LOT Dissertation Series 113.
- Escudero, P., Benders, T., & Lipski, S. (2009). Native, non-native and L2 perceptual cue weighting for Dutch vowels: The case of Dutch, German and Spanish listeners. *Journal of Phonetics*. 37, 452-465. <https://doi:10.1016/j.wocn.2009.07.006>
- Escudero, P., & Chládková, K. (2010). Spanish listeners' perception of American and Southern British English vowels. *The Journal of the Acoustical Society of America*, 128 (5), 254–260. <https://doi.org/10.1121/1.3488794>
- Fabricius, A. (2007). Variation and change in the TRAP and STRUT vowels of RP: A real time comparison of five acoustic data sets. *Journal of the International Phonetic Association*, 37, 293–320.
- Fairbanks, G. (1960). *Voice and articulation drillbook* (2nd ed.). New York: Harper & Row.
- Fant, G. (1970). *Acoustic theory of speech production*. The Hague:Walter de Gruyter.
- Ferguson, S., & Quene, H. (2014). Acoustic correlates of vowel intelligibility in clear and conversational speech for young normal-hearing and elderly hearing-impaired listeners. *The Journal of Acoustical Society of America*, 135, 3570–3584.
- Ferragne, E., & Pellegrino, F. (2010). Formant frequencies of vowels in 13 accents of the British Isles. *Journal of the International Phonetic Association*, 40 (1), 1–34.
- Flege, J. E. (1988). Factors affecting degree of perceived foreign accent in English sentences. *Journal of the Acoustical Society of America*, 84, 70–79.

- Flege, J. E. (1991). The Interlingual Identification of Spanish and English Vowels: Orthographic Evidence. *The Quarterly Journal of Experimental Psychology*, 43 (3), 701-73.
- Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233–277). Baltimore, MD: York Press.
- Flege, J. E. (1999). Age of learning and second language speech. In D. Birdsong (Ed.), *Second language acquisition and the Critical Period Hypothesis* (pp. 101–131). New Jersey: Lawrence Erlbaum Associates Publishers.
- Flege, J. E. (2009). Give input a chance! In T. Piske & M. Young-Scholten (eds.), *Input Matters in SLA*, (pp. 175–190). Bristol: Multilingual Matters.
- Flege, J. E. (2018). It's input that matters most, not age. *Bilingualism: Language and Cognition*, 21 (5), 919-920. [https://doi:10.1017/S136672891800010X](https://doi.org/10.1017/S136672891800010X)
- Flege, J. E. (1999b). *The relation between L2 production and perception*. In the 14th International Congress of Phonetic Sciences, (pp. 1273–1276). San Francisco, California
- Flege, J. E., Bohn, O.-S., & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics*, 25 (4), 437–470.
- Flege, J. E., Munro, M. J., & Skelton, L. (1992). Production of the word-final /t-/d/ contrast by native speakers of English, Mandarin, and Spanish. *Journal of the Acoustical Society of America*, 92, 128–143.
- Flege, J. E., & Liu, S. (2001). The effect of experience on adults' acquisition of a second language. *Studies in Second Language Acquisition*. 23 (4), 527–552.
- Flege, J. E., & Wayland, R. (2019). The role of input in native Spanish late learners' production and perception of English phonetics segments. *Journal of Second Language Studies* 2 (1).
- Flege, J., Aoyama, K., & Bohn, O.-S. (2021). The Revised Speech Learning Model (SLM-r) Applied. In R. Wayland (Ed.), *Second Language Speech Learning: Theoretical and Empirical Progress* (pp. 84-118). Cambridge: Cambridge University Press. [https://doi:10.1017/9781108886901.003](https://doi.org/10.1017/9781108886901.003)
- Fogerty, D., & Humes, L. (2012). The role of vowel and consonant fundamental frequency, envelope, and temporal fine structure cues to the intelligibility of words and sentences. *The Journal of the Acoustical Society of America*, 131 (2), 1490-1501.
- Fraser, H. (2003). Issues in transcription: factors affecting the reliability of transcripts as evidence in legal cases. *International Journal of Speech Language and the Law*, 10 (2), 203-226.

- Fraser, H. (2013). Spoken Word recognition. In C.A. Chapelle, ed. *Encyclopedia of Applied Linguistics*. Oxford: Wiley-Blackwell.
- Fraser, H. (2018). Forensic transcription: How confident false beliefs about language and speech threaten the right to a fair trial in Australia. *Australian Journal of linguistics*, 38 (4), 586-606.
- Fraser, H. (2018b). Real forensic experts should pay more attention to the dangers posed by 'ad hoc' experts. *Australian Journal of Forensic Sciences*, 50 (2), 129-139.
- Fraser, H. (2019). The reliability of voice recognition by ear witnesses: An overview of research findings. *Language and Law*, 6 (2), 1-9.
- French, P., Clermont, F., Harrison, P., & Simpson, S. (2008). *Population data for English spoken in England: A modest first step*. Annual Conference of the International Association for Forensic Phonetics and Acoustics (IAFPA), Lausanne, Switzerland.
- Fullana-Rivera, N., & Mackay, I. (2003). *Production of English sounds by EFL learners: The case of /i/ and /i/*. 15<sup>th</sup> International Congress of Phonetic Sciences, 1525-1528 Barcelona, Spain.
- Gibbon, D., Moore, R., & Winski, R. (1997). *Handbook of standards and resources for spoken language systems*. Berlin: Mouton de Gruyter.
- Gimson, A. (1962). *An introduction to the pronunciation of English*. London: Edward Arnold.
- Gorba, C., & Cebrian, J. (2021). The role of L2 experience in L1 and L2 perception and production of voiceless stops by English learners of Spanish. *Journal of Phonetics*. <https://doi.org/10.1016/j.wocn.2021.101094>
- Grenon, I., Kubota, M., & Sheppard, C. (2019). The creation of a new vowel category by adult learners after adaptive phonetic training. *Journal of Phonetics*, 72, 17-34.
- Guion, S., Flege, J. E., Akahane-Yamada, R., & Pruitt, J. (2000). An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *Journal of the Acoustical Society of America*, 107 (5), 2711-24.
- Haahr, M. (1998). Random.org. Retrieved from <https://www.random.org/company/>
- Haddican, B., Foulkes, P., Hughes, V., & Richards, H. (2013). Interaction of social and linguistic constraints on two vowel changes in northern England. *Language Variation and Change*, 25 (3), 371-403. <https://doi.org/10.1017/S0954394513000197>
- Harrington, J., Kleber, F., & Reubold, U. (2008). Compensation for coarticulation, /u/ fronting, and sound change in standard southern British: An acoustic and perceptual study. *Journal of Acoustic Society of America*, 123 (5), 2825-2835.



- Hattori, K., & Iverson, P. (2009). English /r/-/l/ category assimilation by Japanese adults: Individual differences and the link to identification accuracy. *The Journal of the Acoustical Society of America*, 125, 469–79.
- Hawkins, S., & Midgley, J. (2005). Formant frequencies of RP monophthongs in four age groups of speakers. *Journal of International Phonetic Association*. 35, 183– 199.
- Henton, C. (1983). Changes in the vowels of Received pronunciation. *Journal of Phonetics*, 11, 353-371.
- Hickey, R. (2018). ‘Yes, that’s the best’: Short front vowel lowering in English today: Young people across the anglophone world are changing their pronunciation of vowels according to a change which started in North America. *English Today*, 34(2), 9-16. <https://doi:10.1017/S0266078417000487>
- Hillenbrand, J., Clark, M., & Houde, R. (2000). Some effects of duration on vowel recognition. *Journal of the Acoustical Society of America*, 108 (6), 3013-3022.
- Holt, L., & Lotto, A. (2006). Cue weighting in auditory categorization: Implications for first and second language acquisition. *Journal of the Acoustical Society of America*, 119 (5), 3059–3071.
- Hu, R. (2016). The age factor in second language learning. *Theory and Practice in Language Studies*, 6, (11) 2164-2168.
- Hualde, J. I. (2005). *The Sounds of Spanish*. Cambridge: Cambridge University Press.
- Hualde, J.I., & Prieto, P. (2015). Intonational variation in Spanish: European and American varieties. In S. Frota & P. Prieto (Eds) *Intonation in Romance* (pp 350-391). Oxford. <https://doi.org/10.1093/acprof:oso/9780199685332.003.0010>
- Huber, J. E., Stathopoulos, E. T., Curione, G. M., Ash, T. A., & Johnson, K. (1999). Formants of children, women, and men: The effects of vocal intensity variation. *Journal of Acoustic Society of America*, 106, 1532–1542.
- Hughes, A., Trudgill, P., & Watt, D. (2012). *English Accents and Dialects* (5<sup>th</sup> ed). Hodder Education, Routledge.
- Instituto Cervantes. (2017). *El español en el mundo 2017*. [https://www.cervantes.es/sobre\\_instituto\\_cervantes/prensa/2017/noticias/Presentación-Anuario-2017.htm](https://www.cervantes.es/sobre_instituto_cervantes/prensa/2017/noticias/Presentación-Anuario-2017.htm)
- International Phonetic Association (Ed.). (1999). *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet*. Cambridge: Cambridge University Press.
- Jones, D. (1917). *English pronunciation dictionary*. London: Dent.

- Jones, D. (1926). *English pronunciation dictionary* (3<sup>rd</sup> ed). London: Dent
- Jun, S., & Cowie, I. (1994). Interference of ‘new’ and ‘similar’ vowels in Korean speakers of English. *Ohio State University Working Papers*, 43, 117-130.
- Kim, D., Clayards, M., & Goad, H. (2018). A longitudinal study of individual differences in the acquisition of new vowel contrasts. *Journal of Phonetics*, 67, 1–20.
- Kleber, F., Harrington, J., & Reubold, U. (2011). The relationship between perception and production of coarticulation during a sound change in progress. *Language and Speech*, 55 (3), 383–405.
- Koffi, E., & Lesniak, F. (2019). A longitudinal acoustic phonetic study of English vowels by a Panamanian speaker. *Linguistic Portfolios*, 8, (5), 48-63.
- Koffi, E., & Krause, J. (2020). Speech style variation of vowels in citation forms vs running speech: Intelligibility implications for AI-Enabled devices. *Linguistic Portfolios*, 9, (6), 60-85.
- Kondaurova, M., & Francis, A. (2010). The role of selective attention in the acquisition of English tense and lax vowels by native Spanish listeners: Comparison of three training methods. *Journal of Phonetics*, 38 (4), 569-587.  
[https://doi: 10.1016/j.wocn.2010.08.003](https://doi.org/10.1016/j.wocn.2010.08.003)
- Kuhl, P. (1991). Human adults and human infants show a “perceptual magnet” effect for the prototypes of speech categories, monkeys do not. *Perception & Psychophysics*, 50 (2), 93–107.
- Labov, W. (1972). *Sociolinguistic Patterns*. University of Pennsylvania Press, inc.
- Ladefoged, P. (1996). *Elements of Acoustic Phonetics* (2<sup>nd</sup> ed). The University of Chicago Press.
- Ladefoged, P. (2006). *A Course in Phonetics* (5<sup>th</sup> ed). Boston Massachusetts, USA., Thomson Wadsworth Corporation.
- Ladefoged, P., & Disner, S. (2012). *Vowels and Consonants* (3<sup>rd</sup> ed). Wiley-Blackwell.
- Ladefoged, P., & Johnson, K. (2010). *A Course in Phonetics* (6<sup>th</sup> ed). Boston Massachusetts, USA: Wadsworth, Cengage Learning.
- Lammert, A., Melot, J., Sturim, D., Hannon, D., DeLaura, R., Williamson, J., Ciccarelli, G., & Quatieri, T. (2020). Analysis of phonetic balance in Standard English passages. *Journal of Speech, Language, and Hearing Research*, 63, 917–930.
- Lara, L. (2011). El símbolo, el poder y la lengua. In Silvia Senz & Montserrat Albete (eds.), *El dardo en la Academia. Esencia y vigencia de las academias de la lengua española*, 1, 315–341. Barcelona: Melusina.

- Lee, S., & Cho, M. (2020). The impact of L2-learning experience and target dialect on predicting English vowel identification using Korean vowel categories. *Journal of Phonetics* 82, 1-20. <https://doi.org/10.1016/j.wocn.2020.100983>
- Lee, S., Kang, J., & Nam, H. (2020). Identification of English vowels by non-native listeners: Effects of listeners' experience of the target dialect and talkers' language background. *Second Language Research*. <https://doi.org/10.1177/0267658320965648>
- Leiner, D. (2019). SoSci Survey (Version 3.1.06) [Computer software]. Munich. Retrieved from <https://www.soscisurvey.de>
- Lenneberg, E. (1967). *Biological foundations of language*. New York: Wiley.
- Leow, R. (2007). Input in the L2 classroom: An attentional perspective on receptive practice. In DeKeyser, R. *Practice in a Second Language: Perspective from Applied Linguistics and Cognitive Psychology*. Cambridge Applied Linguistics (pp. 21-50). Cambridge University Press. <https://doi:10.1017/CBO9780511667275.004>
- Leung, K., Jongman, A., Wang, Y., & Sereno, J. A. (2016). Acoustic characteristics of clearly spoken English tense and lax vowels. *The Journal of Acoustical Society of America*, 140, 45-58.
- Levis, J. (2006). Pronunciation and the assessment of spoken language. In R. Hughes (ed.): *Spoken English, TESOL and Applied Linguistics: Challenges for Theory and Practice*. Palgrave Macmillan (pp. 245–70).
- Lipski, J. (2012). *The Handbook of Hispanic Linguistics*. West Sussex, UK: Blackwell Publishing Ltd.
- Llamas, C. (2007). "A place between places": Language and identities in a border town. *Language in Society*, 36 (4), 579-604. <https://doi:10.1017/S0047404507070455>
- Lubowicz, A. (2011). Chain shifts. In Marc van Oostendorp, Colin Ewen, Beth Hume and Keren Rice (Eds.) *Companion to Phonology*, 1-19. Wiley-Blackwell.
- Mairano, P., Bouzon, C., Capliez, M., & De Iacovo, V. (2020). *Acoustic distances, Pillai scores and LDA classification scores as metrics of L2 comprehensibility and nativelikeness*. ICPHS2019,1104-1108. Melbourne, Australia.
- Martínez-Celdrán E., Fernández-Planas, M., & Carrera-Sabaté, J. (2003). Castilian Spanish. *Journal of the International Phonetic Association*, 33, 2. <https://doi:10.1017/S0025100303001373>
- Martínez-Celdrán, E., & Elvira-García, W. (2019). Description of Spanish Vowels and Guidelines for Teaching Them. In: Rao, R. (Ed.), *Key Issues in the Teaching of Spanish Pronunciation*. Oxford: Routledge. pp 17-39. <https://doi:10.4324/9781315666839-2>
- Melià, B. (1992). *La lengua Guaraní del Paraguay*. Madrid: Editorial Mapfre.

- Morrison, G. (2002). Perception of English /i/ and /ɪ/ by Japanese and Spanish listeners: longitudinal results. In G. S. Morrison & L. Zsoldos (Eds.), *North West Linguistics Conference*. Burnaby, BC: Simon Fraser University Linguistics Graduate Student Association.
- Muir, K., Joinson, A., Cotterill, R., & Dewdney, N. (2017). Linguistic Style Accommodation Shapes Impression Formation and Rapport in Computer-Mediated Communication. *Journal of Language and Social Psychology*, 36 (5), 525–548.  
<https://doi.org/10.1177/0261927X17701327>
- Munro, M., & Derwing, T. (1995). Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech, *Language and Speech*, 38, 289–306.
- Munro, M., & Derwing, T. (2008). Segmental acquisition in adult ESL learners: A longitudinal study of vowel production. *Language Learning*, 58 (3), 479-502.
- Nittrouer, S., & Lowenstein, J. H. (2010). Learning to perceptually organize speech signals in native fashion. *Journal of the Acoustical Society of America*, 127 (3), 1624–1635.
- Nycz, J., & Hall-Lew, L. (2013). *Best practices in measuring vowel merger*. Proceedings on Meetings on Acoustics, 20. <https://doi.org/10.1121/1.4894063>
- Ogden, R. (2009). *An Introduction to English Phonetics*. Edinburgh: Edinburgh University Press.
- O'Rourke, E. (2005). *Intonation and language contact: A case study of two varieties of Peruvian Spanish*. [Doctoral dissertation, University of Illinois at Urbana-Champaign]. ProQuest Dissertation Publishing.
- Patel, R., Connaghan, K., Franco D., Edsall E., Forgit, D., Olsen, L., Ramage, L., Tyler, E., & Russell, S. (2012). The Caterpillar: A Novel Reading Passage for Assessment of Motor Speech Disorders, *American Journal of Speech-Language Pathology and American Speech-Language-Hearing Association*.
- Pharies, D. (2007). *A brief history of Spanish language*. The University of Chicago Press, Ltd., London.
- Piske, T., MacKay, I., & Flege, J. E. (2001). Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics*, 29 (2), 191–215.  
<https://doi.org/10.1006/jpho.2001.0134>
- Polka, L. (1991). Cross-language speech perception in adults: Phonemic, phonetic, and acoustic contributions. *Journal of Acoustical Society of America*, 89, 2961-2977.
- Polka, L. (1995). Linguistic influences in adult perception of non-native vowel contrasts. *Journal of the Acoustical Society of America*, 97, 1286-1296.

- Pols, L., Van der Kamp, L., & Plomp, R. (1969). Perceptual and physical space of vowel sounds. *Journal of the Acoustical Society of America* 46, 458- 467.
- Real Academia Española. (2019). <https://www.rae.es/la-institucion>
- Roach, P. (2004). British English Received Pronunciation. *Journal of the International Phonetic Association*, 34, 2. <https://doi:10.1017/S0025100304001768>
- Roach, P. (2009). *English phonetics and phonology. A practical course* (2<sup>nd</sup> ed.). Cambridge: Cambridge University Press.
- Rato, A., & Carlet, A. (2020). Second language perception of English vowels by Portuguese learners: The effect of stimulus type. *Journal ILHA Do Desterro*, 75 (3), 205-226. <http://dx.doi.org/10.5007/2175-8026.2020v73n3p205>
- Ronquest, R. (2018). Vowels. In K. Geeslin (Ed.), *The Cambridge Handbook of Spanish Linguistics*. Cambridge Handbooks in Language and Linguistics (pp.145-164). Cambridge University Press. <https://doi:10.1017/9781316779194.008>
- Sala, M., & Posner, R. (2021). *Spanish language*. Encyclopedia Britannica. <https://www.britannica.com/topic/Spanish-language>
- Salcedo, C. (2010). The Phonological System of Spanish. *Revista de Lingüística y lenguas aplicadas*, 5, 195-205.
- Sancier, M., & Fowler, C. (1997). Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of Phonetics*, 25, 421–438.
- Selinger, H. W., & Shohamy, E. (1989). *Second Language Research Methods*. Oxford: Oxford University Press.
- Shattuck-Hufnagel, S., & Turk, A. (1996). A Prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research*, 25 (2), 193-247. <https://doi:10.1007/BF01708572>
- Smiljanić, R., & A.R. Bradlow. (2005). Production and perception of clear speech in Croatian and English. *The Journal of the Acoustical Society of America*, 118, 1677-1688.
- Smiljanić, R., & A.R. Bradlow. (2008). Stability of temporal contrasts across speaking styles in English and Croatian. *Journal of Phonetics*, 36 (1), 91-113.
- Smith, B., Johnson, E., & Hayes-Harb, R. (2019). ESL learners' intra-speaker variability in producing American English tense and lax vowels. *Journal of Second Language Pronunciation*, 5 (1), 139–164. <https://doi.org/10.1075/jslp.15050.smi>
- Song, J.Y., & Eckman, F. (2019). Covert contrasts in the acquisition of English high front vowels by native speakers of Korean, Portuguese, and Spanish. *Language Acquisition*, 26, 436-456. <https://doi:10.1080/10489223.2019.1593415>

- Sóskuthy, M., Foulkes, P., Haddican, W., Hay, J., & Hughes, V. (2015). *Word-level distributions and structural factors codetermine GOOSE fronting*. Proceedings of ICPhS 18, Glasgow, UK: the University of Glasgow
- Strange, W., Weber, A., Levy, E., Shafiro, V., Hisagi, M., & Nishi, K. (2007). Acoustic variability within and across German, French, and American English vowels: Phonetic context effects. *Journal of the Acoustical Society of America*, 122, 1111-1129.
- Strycharczuk, P., & Scobbie, J. (2017). Fronting of Southern British English high-back vowels in articulation and acoustics. *The Journal of the Acoustical Society of America*, 142-322. <https://doi:10.1121/1.4991010>
- Taylor, M. (2022). La Diversidad Lingüística Durante y Después del Franquismo en España. The Review. *A Journal of Undergraduate Student Research* 23.
- Thomas, E., & Kendall, T. (2007). NORM: The vowel normalization and plotting suite. [Online Program]. Retrieved from: <http://ncslaap.lib.ncsu.edu/tools/norm/>
- Upton, C., Kretzschmar, W. & Konopka, R. (2001). *Oxford Dictionary of Pronunciation for Current English*. Oxford: Oxford University Press.
- Vann, R. (1999). Language exposure in Catalonia: An example of indoctrinating linguistic ideology. *Word*, 50 (2), 191-209. <https://doi:10.1080/00437956.1999.11432488>
- Van Leussen, J-W., Williams, D., & Escudero, P. (2011). *Acoustic properties of Dutch steady-state vowels: Contextual effects and a comparison with previous studies*, in W. Lee and E. Zee (Eds.), 17th International Congress of Phonetic Sciences (pp. 1194-1197).
- Van Riper, C. (1963). *Speech correction: Principles and methods* (4th ed.). Englewood Cliffs, NJ: Prentice-Hall.
- Wang, X., & Munro, M. J. (1999). *The perception of English tense-lax vowel pairs by native Mandarin speakers: The effect of training on attention to temporal and spectral cues*. 14<sup>th</sup> International Congress of Phonetic Sciences, 3, 125-128.
- Watt, D., & Fabricius, A. (2002). Evaluation of a technique for improving the mapping of multiple speakers' vowel spaces in the F1~ F2 plane. *Leeds working papers in linguistics and phonetics*, 9 (9), 159- 173.
- Watt, D., Fabricius, A., & Kendall, T. (2010). More on Vowels: Plotting and Normalization. In M. Di Paolo, & M. Yaeger-Dror (Eds.), *Sociophonetics: a student's guide* (pp. 107-118). Routledge.
- Wells, J. (1962). A study of the formants of the pure vowels of British English. [M.A. thesis University of London]. Online.
- Wells, J. (1982). *Accents of English*. Cambridge: Cambridge University Press.

- Wells, J. (2001). IPA transcription systems for English.  
<https://www.phon.ucl.ac.uk/home/wells/ipa-english-uni.htm>
- Wiik, K. (1965). *Finnish and English vowels*. Turku: University of Turku
- Wright, R., & Nichols, D. (2015). Measuring Vowel Formants.  
<https://depts.washington.edu/phonlab/resources/measuring-formants.pdf>
- Yazawa, K., Whang, J., Kondo, M., & Escudero, P. (2020). Language-dependent cue weighting: An investigation of perception modes in L2 learning. *Second Language Research*, 36 (4), 557–581.
- Zahorian, S., & Jagharghi, A. (1993). Spectral-shape features versus formants as acoustic correlates for vowels, *The Journal of Acoustical Society of America*, 94, 1966–1982.

## Appendices

### Appendix A. Information sheet

**Department of Language and Linguistic Science**

Heslington, York, YO10 5DD, UK

#### INFORMATION SHEET

PLEASE KEEP THIS INFORMATION SHEET AND A SIGNED COPY OF THE  
CONSENT FORM FOR YOUR RECORDS

*You are invited to take part in a research study. Before you decide whether to participate it is important for you to understand why the research is being done and what it will involve. Please take the time to read the following information carefully. If there is anything you do not understand, or if you want more information, please ask the researcher.*

Title of study

**Developmental perception and production of English Speech sounds by EFL Spanish speakers**

Researcher: María Gabriela Valenzuela Farías

What is the research about?

The aim of this study is to analyse speech sounds produced by native Spanish speakers. It seeks to establish whether and how exposure, as a result of immersion in the target language, modifies the perception/production of speech sounds. It also seeks to determine which are the most important factors affecting the accuracy of production and perception of problematic English sounds.

Who is carrying out the research?

The research will be carried out by María Gabriela Valenzuela Farías.

Why have you been chosen to participate?

You have been invited to participate because you are starting your studies at the University of York and you are a native speaker of Spanish.



### What does the study involve?

The study will take place in the phonetics laboratory in the Linguistics department. It will involve you attending three times over the academic year. The first session will be in November 2018, then February 2019 and finally in August 2019. Each session should last approximately 20 minutes. It will involve 2 parts. The first part consists of reading two short paragraphs aloud, describing a picture and reading a list of isolated words in English. The second part consists of listening to a word in English and at the same time looking at a screen with two words for selecting the one it has been played.

The data will be collected by the researcher. She and the participant will be in the room with the technical supervisor of the Linguistic Dept. while the data is being collected.

### Do I have to take part?

You do not have to take part in the study. If you do decide to take part you will be given this information sheet to keep and will be asked to sign two copies of the consent form (one copy is for you to keep). If you decide to take part you will still be free to withdraw without giving a reason, even during the session itself. If you withdraw from the study, we will destroy your data and will not use it in any way.

### What are the possible risks of taking part?

The risks (if any) associated with this study are not greater than the risk of talking or reading aloud in everyday life.

### Are there any benefits to participating?

By participating in this study, you will help to build new knowledge in the field of phonetics, and will contribute to research related to EFL speakers with Spanish as their native language.

### What will happen to the data I provide?

The data you provide will be used alongside the data of other participants to be analysed by measuring production and perception of English speech. Your data will be stored securely in the University of York, Department of Language and Linguistic Science for a minimum period of five years, after that I will review it and decide if it should be destroyed or kept for more time.

### What about confidentiality?

Your identity will be kept strictly confidential. No real names will be used in any presentations, publications or in my dissertation.

Will I know the results?

Only group results will be given – if requested. No individual feedback will be given to participants.

*This study has been reviewed and approved by the Departmental Ethics Committee of the Department of Language and Linguistic Science at the University of York. If you have any questions regarding this, you can contact the chair of the L&LS Ethics Committee.*

If you have further questions regarding this study, please feel free to contact:

M.Gabriela Valenzuela Farías  
Department of Language and Linguistic Science  
University of York, Heslington, York, YO10 5DD  
**email:** mgvf500@york.ac.uk

Prof. Peter French  
Department of Language and Linguistic Science  
Phone:44(0)1904 322649  
Email: peter.french@york.ac.uk

## Appendix B. Consent form

### Developmental perception and production of English Speech Sounds by EFL Spanish speakers

Researcher: María Gabriela Valenzuela Farías.

#### Consent form

This form is for you to state whether or not you agree to take part in the study. Please read and answer every question. If there is anything you do not understand, or if you want more information, please ask the researcher.

Have you read and understood the information leaflet about the study? Yes  No

Have you had an opportunity to ask questions about the study and have these been answered satisfactorily? Yes  No

Do you understand that the information you provide will be held in confidence by the researcher, and your name or identifying information about you will not be mentioned in any publication? Yes  No

Do you understand that you may withdraw from the study at any time before the end of the data collection session without giving any reason, and that in such a case all your data will be destroyed? Yes  No

Do you understand that the information you provide may be kept after the duration of the current project, to be used in future research on language? Yes  No

Do you agree to take part in the study? Yes  No

Your name (in BLOCK letters): \_\_\_\_\_

Your signature: \_\_\_\_\_

Researcher's name: María Gabriela Valenzuela Farías.

Date: \_\_\_\_\_

One copy to be retained by the researcher, one copy to be kept by the participant.

## Appendix C. Reading materials

### 1. Passages

#### The North Wind and the Sun

The North Wind and the Sun were disputing which was the stronger, when a traveller came along wrapped in a warm cloak. They agreed that the one who first succeeded in making the traveller take his cloak off should be considered stronger than the other. Then the North Wind blew as hard as he could, but the more he blew the more closely did the traveller fold his cloak around him, and at last the North Wind gave up the attempt. Then the Sun shone out warmly, and immediately the traveller took off his cloak. And so the North Wind was obliged to confess that the Sun was the stronger of the two.

## The Grandfather Passage

You wished to know all about my grandfather. Well, he is nearly ninety-three years old; he dresses himself in an ancient black frock coat, usually minus several buttons; yet he still thinks as swiftly as ever. A long, flowing beard clings to his chin, giving those who observe him a pronounced feeling of the utmost respect. When he speaks, his voice is just a bit cracked and quivers a trifle. Twice each day he plays skillfully and with zest upon our small organ. Except in the winter when the ooze or snow or ice prevents, he slowly takes a short walk in the open air each day. We have often urged him to walk more and smoke less, but he always answers, "Banana oil!" Grandfather likes to be modern in his language.

## **The Boy who Cried Wolf**

There was once a poor shepherd boy who used to watch his flocks in the fields next to a dark forest near the foot of a mountain. One hot afternoon, he thought up a good plan to get some company for himself and also have a little fun. Raising his fist in the air, he ran down to the village shouting 'Wolf, Wolf.' As soon as they heard him, the villagers all rushed from their homes, full of concern for his safety, and two of his cousins even stayed with him for a short while. This gave the boy so much pleasure that a few days later he tried exactly the same trick again, and once more he was successful. However, not long after, a wolf that had just escaped from the zoo was looking for a change from its usual diet of chicken and duck. So, overcoming its fear of being shot, it actually did come out from the forest and began to threaten the sheep. Racing down to the village, the boy of course cried out even louder than before. Unfortunately, as all the villagers were convinced that he was trying to fool them a third time, they told him, 'Go away and don't bother us again.' And so, the wolf had a feast.

## **Caterpillar Passage**

Do you like amusement parks? Well, I sure do. To amuse myself, I went twice last spring. My most memorable moment was riding on the Caterpillar, which is a gigantic roller coaster high above the ground. When I saw how high the Caterpillar rose into the bright blue sky I knew it was for me. After waiting in line for thirty minutes, I made it to the front where the man measured my height to see if I was tall enough. I gave the man my coins, asked for change, and jumped on the cart. Tick, tick, tick, the Caterpillar climbed slowly up the tracks. It went so high I could see the parking lot. Boy was I scared! I thought to myself, "There's no turning back now." People were so scared they screamed as we swiftly zoomed fast, fast, and faster along the tracks. As quickly as it started, the Caterpillar came to a stop. Unfortunately, it was time to pack the car and drive home. That night I dreamt of the wild ride on the Caterpillar. Taking a trip to the amusement park and riding on the Caterpillar was my most memorable moment ever!

### The Story of Arthur the Rat

There was once a young rat named Arthur who would never take the trouble to make up his mind. Whenever his friends asked him if he would like to go out with them, he would only answer, "I don't know." He wouldn't say "Yes" and he wouldn't say "No" either. He could never learn to make a choice. His Aunt Helen said to him "No-one will ever care for you if you carry on like this. You have no more mind than a blade of grass." Arthur looked wise but said nothing. One rainy day the rats heard a great noise in the loft where they lived. The pine rafters were all rotten, and at last one of the joists had given way and fallen to the ground. The walls shook and the rats' hair stood on end with fear and horror. "This won't do," said the old rat who was chief, "I'll send out scouts to search for a new home." Three hours later the seven scouts came back and said, "We've found a stone house which is just what we wanted. There's room and good food for us all. There's a kindly horse named Nelly, a cow, a calf and a garden with an elm tree." Just then the old rat caught sight of young Arthur. "Are you coming with us?" he asked. "I don't know," Arthur sighed, "The roof may not come down just yet." "Well," said the old rat angrily, "We can't wait all day for you to make up your mind. Right about face! March!" And they went off. Arthur stood and watched the other rats hurry away. The idea of an immediate decision was too much for him. "I'll go back to my hole for a bit," he said to himself, "just to make up my mind." That night there was a great crash that shook the earth, and down came the whole roof. Next day some men rode up and looked at the ruins. One of them moved a board, and under it they saw a young rat lying on his side, quite dead, half in and half out of his hole.



### The Rainbow Passage

When the sunlight strikes raindrops in the air, they act like a prism and form a rainbow. The rainbow is a division of white light into many beautiful colours. These take the shape of a long round arch, with its path high above, and its two ends apparently beyond the horizon. There is, according to legend, a boiling pot of gold at one end. People look, but no one ever finds it. When a man looks for something beyond his reach, his friends say he is looking for the pot of gold at the end of the rainbow. Throughout the centuries men have explained the rainbow in various ways. Some have accepted it as a miracle without physical explanation. To the Hebrews it was a token that there would be no more universal floods. The Greeks used to imagine that it was a sign from the gods to foretell war or heavy rain. The Norsemen considered the rainbow as a bridge over which the gods passed from earth to their home in the sky. Other men have tried to explain the phenomenon physically. Aristotle thought that the rainbow was caused by reflection of the sun's rays by the rain. Since then physicists have found that it is not reflection, but refraction by the raindrops which causes the rainbow. Many complicated ideas about the rainbow have been formed. The difference in the rainbow depends considerably upon the size of the water drops, and the width of the coloured band increases as the size of the drops increases. The actual primary rainbow observed is said to be the effect of superposition of a number of bows. If the red of the second bow falls upon the green of the first, the result is to give a bow with an abnormally wide yellow band, since red and green lights when mixed form yellow. This is a very common type of bow, one showing mainly red and yellow, with little or no green and blue.

## 2. Word lists

List 1. Read the following words:

- |          |          |          |
|----------|----------|----------|
| 1. pool  | 21. bust | 42. pair |
| 2. seat  | 22. toy  | 43. suit |
| 3. beat  | 23. pail | 44. pert |
| 4. pit   | 24. best | 45. pot  |
| 5. soft  | 25. putt | 46. peat |
| 6. bit   | 26. post | 47. sent |
| 7. cure  | 27. beer | 48. sort |
| 8. pull  | 28. pier | 49. salt |
| 9. mind  | 29. but  | 50. palm |
| 10. sat  | 30. soot | 51. Bert |
| 11. boot | 31. cow  | 52. tour |
| 12. bird | 32. set  | 53. put  |
| 13. past | 33. pear | 54. fail |
| 14. bat  | 34. part | 55. find |
| 15. pall | 35. bolt | 56. pat  |
| 16. port | 36. how  | 57. pint |
| 17. boat | 37. can  | 58. for  |
| 18. sit  | 38. peer | 59. joy  |
| 19. bear | 39. boy  | 60. now  |
| 20. bait | 40. pure |          |
|          | 41. the  |          |

List 2. Read the following words

- |          |           |             |
|----------|-----------|-------------|
| 1. dust  | 22. deer  | 43. comma   |
| 2. found | 23. light | 44. near    |
| 3. hope  | 24. foul  | 45. feather |
| 4. peal  | 25. sift  | 46. coin    |
| 5. font  | 26. beet  | 47. fair    |
| 6. pant  | 27. dour  | 48. pal     |
| 7. daft  | 28. full  | 49. dirt    |
| 8. peat  | 29. hail  | 50. host    |
| 9. fight | 30. belt  | 51. dart    |
| 10. poor | 31. hurt  | 52. foil    |
| 11. moor | 32. fast  | 53. house   |
| 12. pill | 33. burnt | 54. there   |
| 13. food | 34. foot  | 55. poll    |
| 14. pest | 35. pact  | 56. hair    |
| 15. fort | 36. shut  | 57. point   |
| 16. buy  | 37. dot   | 58. hall    |
| 17. fist | 38. fool  | 59. China   |
| 18. hoot | 39. paid  | 60. hut     |
| 19. push | 40. fear  |             |
| 20. shot | 41. bent  |             |
| 21. gait | 42. fall  |             |

List 3. Read the following words

	21. bump	42. fewer
1. bowl	22. took	43. dad
2. coil	23. boss	44. never
3. bag	24. park	45. doctor
4. dead	25. bin	46. surf
5. bush	26. soil	47. down
6. gear	27. beg	48. soup
7. cheer	28. piece	49. dance
8. dark	29. born	50. lure
9. soul	30. dare	51. doll
10. day	31. hear	52. tiger
11. bug	32. door	53. date
12. deal	33. bike	54. bus
13. where	34. shame	55. chair
14. board	35. dock	56. dean
15. pound	36. serve	57. shoe
16. bite	37. dour	58. bid
17. pole	38. bill	59. join
18. shook	39. dent	60. doubt
19. purse	40. bye	
20. duke	41. dash	

3. Picture description task. Time 1



Time 2



Time 3



## Appendix D

### Language Background Questionnaire

#### Language Background Questionnaire (This information will remain confidential)

**A. Personal information**

---

1. ID. \_\_\_\_\_

**B. Education and language use**

---

1. Which of the four English skills do you feel more confident with? (listening, speaking, reading, writing)

\_\_\_\_\_

2. Which skills are/is most difficult for you? (listening, speaking, reading, writing)

\_\_\_\_\_

3. How motivated do you feel for speaking English?

1      2      3      4      5      6      7      8      9      10

Low motivation High motivation

4. How important is to achieve native-like pronunciation for you?

1      2      3      4      5      6      7      8      9      10

Not important Extremely important

5. How important is for you to maintain your Spanish accent in the English you speak?

1      2      3      4      5      6      7      8      9      10

Not important Extremely important

6. What variety of English do you like and hope to achieve? (British / American English)

\_\_\_\_\_

7. At the University of York how many hrs a week approximately do you spend in ...

	None 0	1-3hrs	4-6hrs	More than 6
Seminar classes				
Supervisor meeting				
Tutoring time				





21. During the week, how many hours do you spend at the University?

0 1 2 3 4 5 6 7 8 9 10 or more hrs.

22. How many hours do you speak Spanish during the day?

0 1 2 3 4 5 6 7 8 9 10 or more hrs.

23. Do you work on/off campus? (paid job in the UK) \_\_\_\_\_ If "yes"

How many hours do you use or hear English language at work?

0 1 2 3 4 5 6 7 8 9 10 or more hrs.

What is the main accent that you are exposed to during work time? \_\_\_\_\_

24. How important is for you to develop accurate listening and speaking English skills?

1 2 3 4 5 6 7 8 9 10

Not important

Extremely important

25. How important is for you to develop accurate reading and writing English skills?

1 2 3 4 5 6 7 8 9 10

Not important

Extremely important

26. Do you consider that you speak more English or Spanish during a normal day?

\_\_\_\_\_

27. Do you consider that you listen to more English (from native speaker) or Spanish during a normal day?

\_\_\_\_\_

28. Do you speak more English with native speakers or international people?

\_\_\_\_\_

29. If there is/are any other aspect of your language use that it was not mentioned, and you would like to specify, please comment below.

**Appendix E**  
**Comparative values**

*Comparative Spanish group results with RP values*

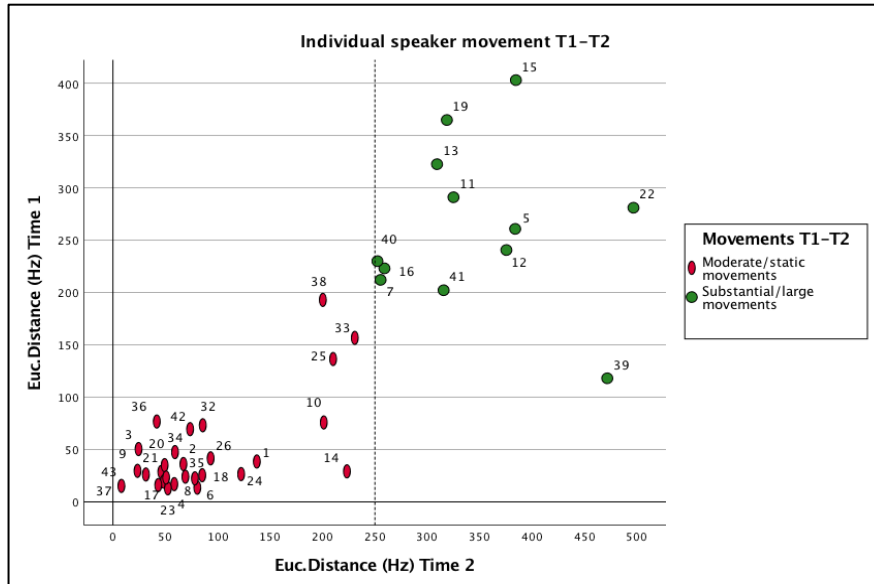
<b>English</b>	<b>F1</b>	<b>F2</b>	<b>Dist.</b>	<b>Spanish</b>	<b>F1</b>	<b>F2</b>	<b>Dist.</b>
/i:/	280	2221	/i:/- /ɪ/	/i:/	343	2477	/i:/- /ɪ/
/ɪ/	380	1960	250	/ɪ/	407	2250	235
/u:/	302	1130	/u:/- /ʊ/	/u:/	389	1230	/u:/- /ʊ/
/ʊ/	414	1051	110	/ʊ/	451	1160	93
/e/	560	1797	/ɪ/-/e/	/e/	625	1922	/ɪ/-/e/
			250				394

## Appendix F

### Individual speakers' trajectories: Word list

#### F1. /i:/and /ɪ/

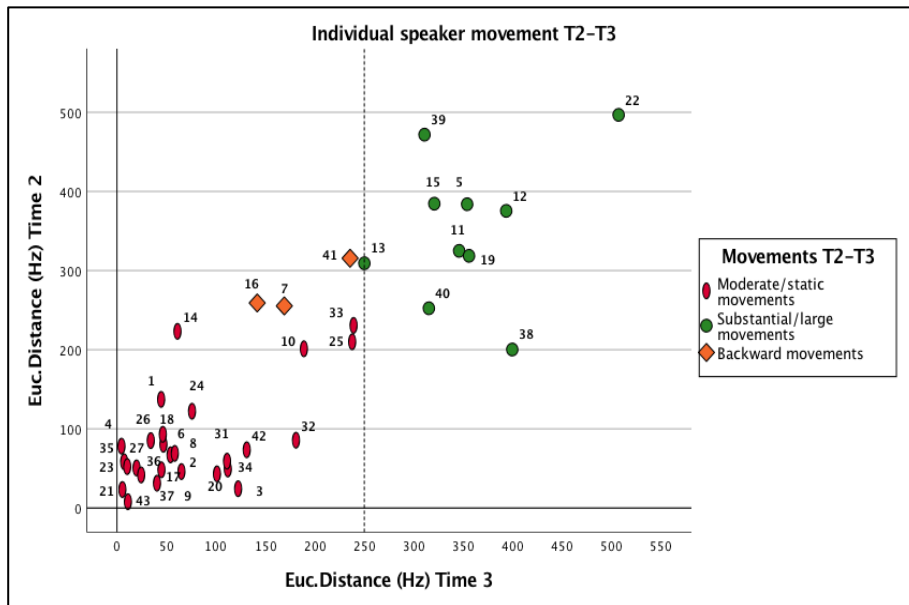
1) Vowel distance variations between /i:/ and /ɪ/ from Time 1-2



#### Summary Table

<i>Summary T1-T2</i>	N	%
a) Moderate/static movement (low performers)	28	70%
b) Substantial/large movement (high performers)	12	30%

2) Vowel distance variations between /i:/ and /i/ 2-3



Summary Table

Summary T2-T3	N	%
a) Moderate/static movement (low performers)	27	67.5%
b) Substantial/large movement (high performers)	10	25%
c) Backward movement (low performers)	3	7.5%

## F2. F1 and F2 results

Overall results showing the number of speakers producing higher or lower F1 values from Time 1-2

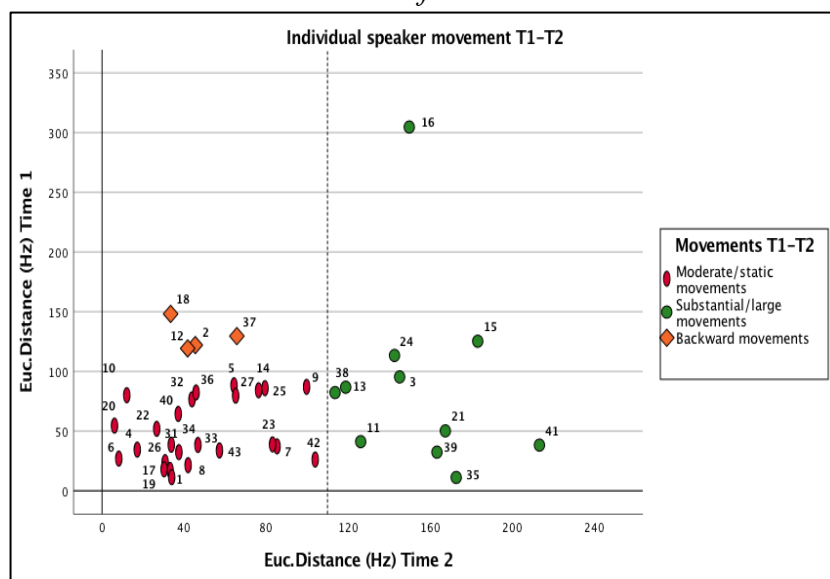
Vowel	N° sp	F1	Vowel	N°sp	F1
/i:/	20	Higher values (avg: 376Hz)	/ɪ/	30	Higher values (avg: 426 Hz)
	19	Lower values (avg: 324Hz)		10	Lower values (avg: 424Hz)

Overall results showing the number of speakers producing higher or lower F2 values Time 1-2

Vowel	N° sp	F2	Vowel	N°sp	F2
/i:/	18	Increased values (fronted) (avg: 2482Hz)	/ɪ/	8	Increased values (fronted) (avg: 1855Hz)
	22	Decreased values (central) (avg: 2422Hz)		32	Decreased values (central-back) (avg: 2138Hz)

## F.3 /u:/ and /ʊ/

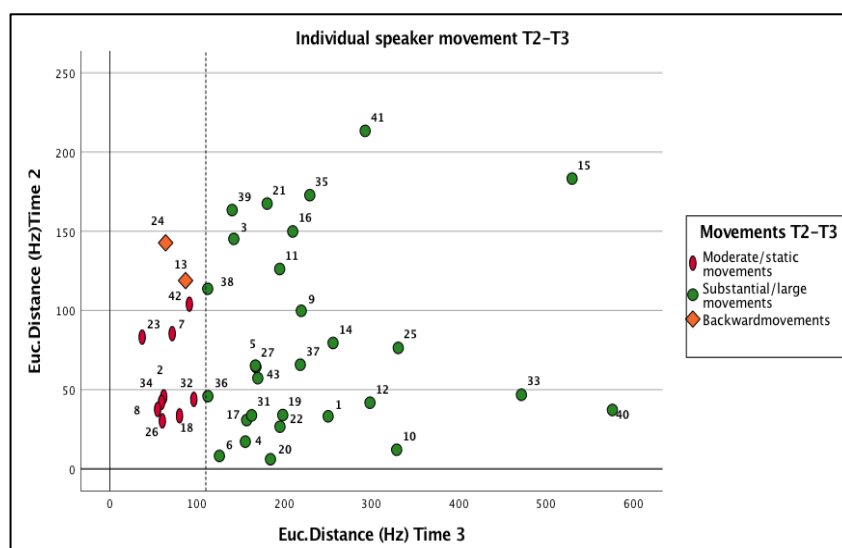
Vowel distance variations between /u:/ and /ʊ/ from Time 1-2



Summary Table

Summary T1-T2	N	%
a) Moderate/static movement (low performers)	25	62.5%
b) Substantial/large movement (high performers)	11	27.5%
c) Backward movement (low performers)	4	10%

Vowel distance variations between /u:/ and /ʊ/ from Times 2-3



Summary Table

Summary T2-T3	N	%
a) Moderate/static movement (low performers)	9	22%
b) Substantial/large movement (high performers)	29	72.5%
c) Backward movement (low performers)	2	5%

#### F4. F1-F2 results

*Overall results showing the number of speakers producing higher or lower F1 values from Time 1-2*

<b>Vowel</b>	<b>N° sp</b>	<b>F1</b>	<b>Vowel</b>	<b>N°sp</b>	<b>F1</b>
/u:/	23	Higher values (avg: 422Hz)	/ʊ/	22	Higher values (avg: 463Hz)
	17	Lower values (avg: 372Hz)		18	Lower values (avg: 415Hz)

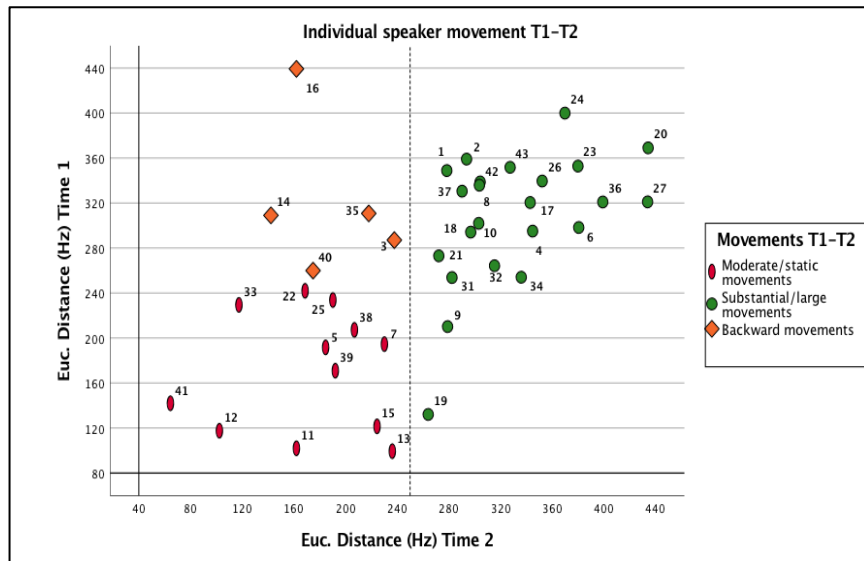
*Overall results showing the number of speakers producing higher or lower F2 values from Time 1-2*

<b>Vowel</b>	<b>N° sp</b>	<b>F2</b>	<b>Vowel</b>	<b>N°sp</b>	<b>F2</b>
/u:/	2	Increased values (fronted) (aver: 1007Hz)	/ʊ/	12	Increased values (fronted) (aver:1069Hz)
	38	Decreased (retracted) (aver: 979Hz)		28	Decreased (retracted) (aver: 1019Hz)



**F5. /ɪ/ and /e/**

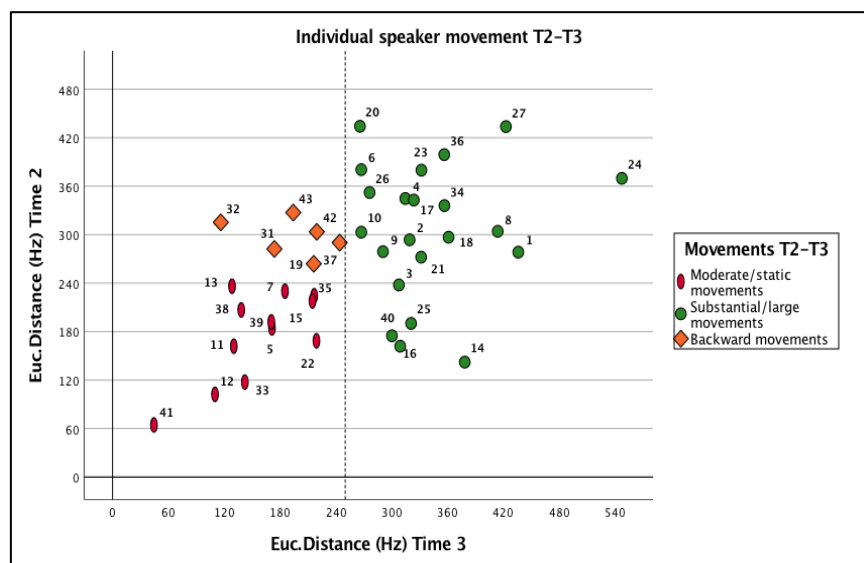
*Vowel distance variations between /ɪ/ and /e/ from Time 1-2*



*Summary Table*

<i>Summary T1-T2</i>	<b>N</b>	<b>%</b>
a) Moderate/static movements (low performers)	12	30%
b) Substantial/large movements (high performers)	24	60 %
c) Backward movement (low performers)	4	10%

*Vowel distance variations between /ɪ/ and /e/ from Time 2-3*



*Summary Table*

<i>Summary T2-T3</i>	<b>N</b>	<b>%</b>
a) Moderate/static movements (low performers)	11	27.5%
b) Substantial/large movements (high performers)	23	57.5%
c) Backward movement (low performers)	6	15%

**F6 F1-F2**

*Overall results showing the number of speakers producing higher or lower F1 values from Time 1-2*

<b>Vowel</b>	<b>N° sp</b>	<b>F1</b>	<b>Vowel</b>	<b>N°sp</b>	<b>F1</b>
/i/	30	Higher values (avg: 426 Hz)	/e/	20	Higher values (avg: 676Hz)
	10	Lower values (avg: 424Hz)		20	Lower values (avg: 588Hz)

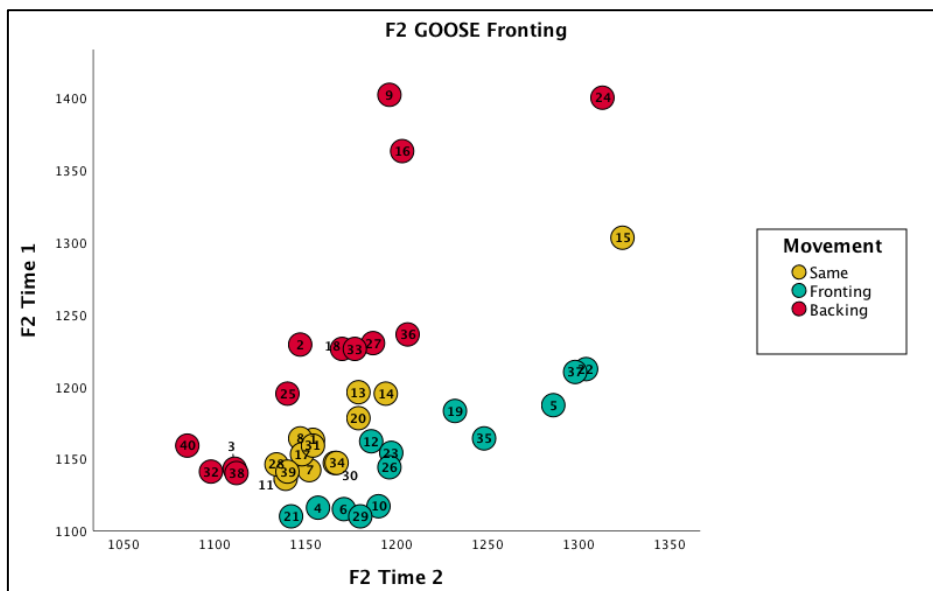
*Overall results showing the number of speakers producing higher or lower F2 values from Time 1-2*

<b>Vowel</b>	<b>N° sp</b>	<b>F2</b>	<b>Vowel</b>	<b>N°sp</b>	<b>F2</b>
/i/	8	Increased values (fronted) (aver: 1855Hz)	/e/	13	Increased values (fronted) (aver:1942Hz)
	32	Decreased values (central-back) (aver: 2138Hz)		27	Decreased values (retracted) (aver: 1836)

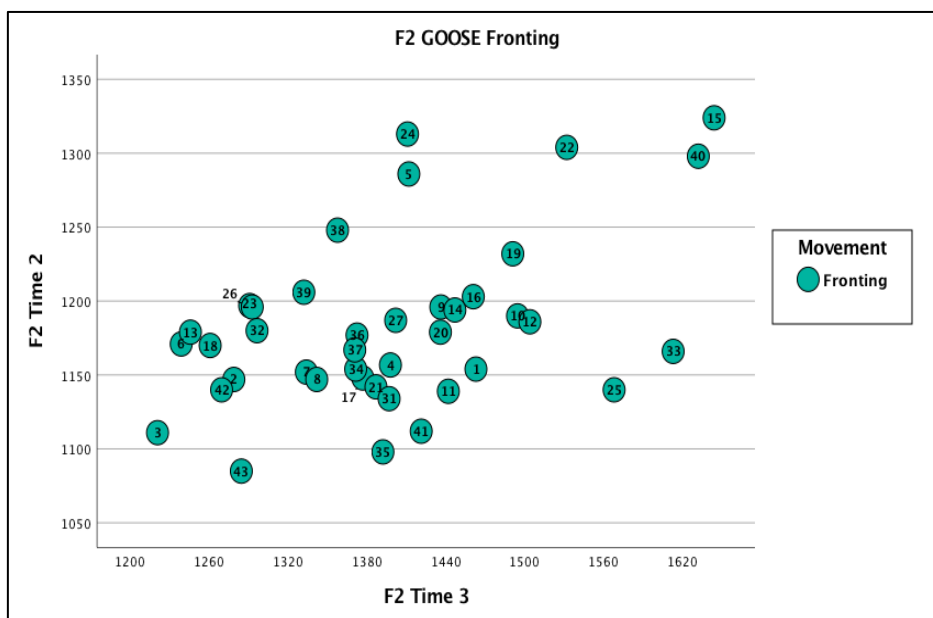
## Appendix G /u:/-/ʊ/ Fronting

### G1 /u:/

Changes in the F2 values for /u:/ between *T1-T2*

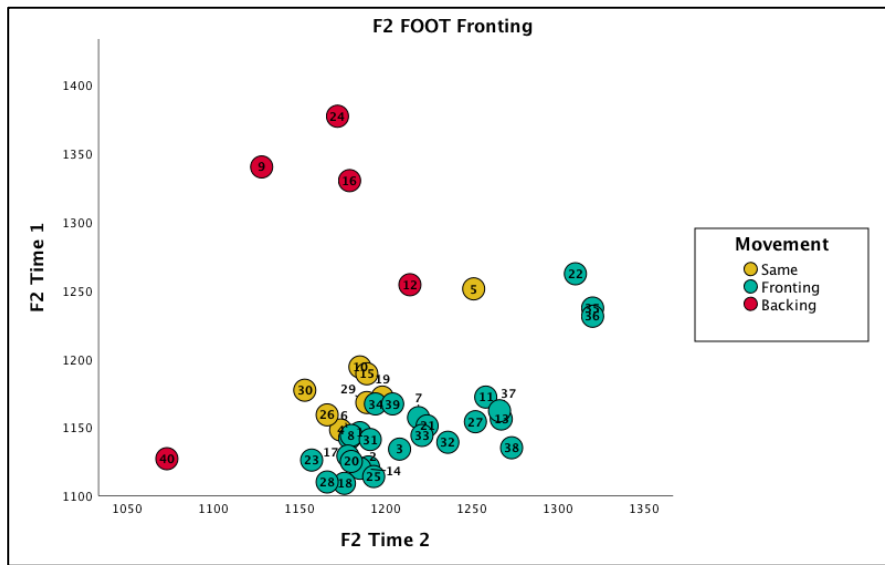


Changes in the F2 values for /u:/ between *T2-T3*

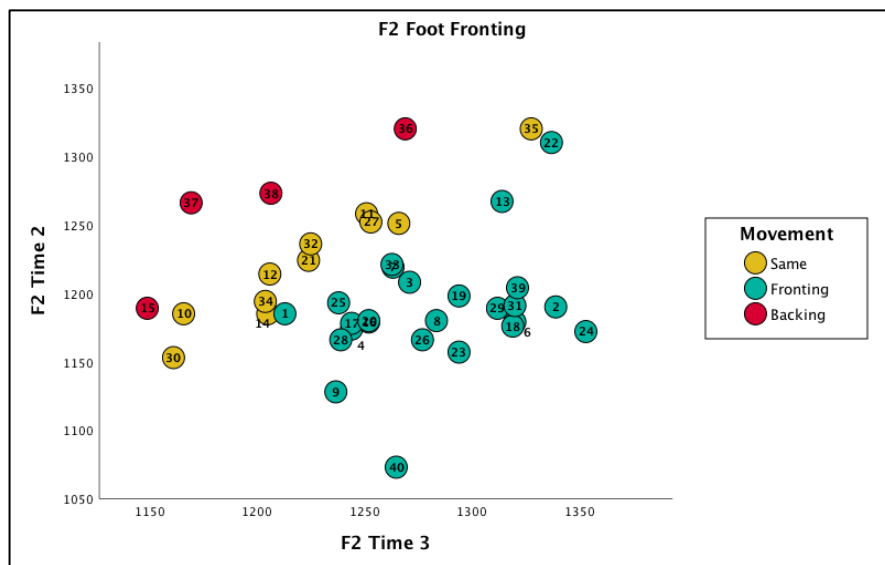


## G2 /ʊ/

Changes in the F2 values for /ʊ/ between T1-T2



Changes in the F2 values for /ʊ/ between T2-T3



## Appendix H

### Perception test participant information

---

#### Perception test participant information

---

##### a) Educational level:

---

The educational attainment for the participants recruited was divided according to the highest level they have achieved, the division was made in the following categories: a) doctoral degree (13), b) master's degree or equivalent (54), c) bachelor's level or equivalent (56), d) A-level or equivalent (16), e) GCSE or equivalent (7), and f) no qualifications (2).

##### b) Language background:

---

All participants were native British English speakers from different regions of England. No one reported speaking Spanish as a foreign language, and only four people stated that they had an elementary level of familiarity with it.

The majority of the respondents (134) did not have a degree in a foreign language. Only 9 out of 148 reported holding a degree in linguistics.

##### c) Gender and Age:

---

In terms of gender this group of participants identified themselves as follows: female (107), male (38), and non-binary (1). Regarding age, all contributors were adults, and their age groups are broken down into four bands: a) 18-29 yr. (19), b) 30-44 yr. (51), c) 45-60 yr. (34) c) and 60+ yr. (34).

---

## Appendix I

### Perception Test, Information electronic sheet

#### PARTICIPANT INFORMATION

##### **What is this research about?**

This task is part of my PhD research. My research investigates speech produced by non-native English speakers. It seeks to establish whether and how exposure, as a result of immersion in the English language, modifies non-native speech. This part of the study aims to collect judgements from native speakers of English on utterances produced by non-native speakers of English.

##### **What does taking part involve?**

If you agree to participate in this task, you will be asked to listen to different words and identify the word you heard. Also, you will be asked some questions regarding the word selection, and general background information. The task will take approximately 15-20 minutes to complete.

##### **Do I have to take part?**

You do not have to take part in the study. If you do decide to take part you will need to agree in the consent form.

##### **What are the possible risks of taking part?**

There are no risks associated to this study.

##### **What will happen to the data I provide?**

Your data will be stored securely for a minimum period of five years, after that I will review it and decide if it should be destroyed or kept longer.

##### **Will I know the results?**

No individual feedback will be given to participants.

##### **On what basis will you process my data?**

Under the General Data Protection Regulation (GDPR), the University has to identify a legal basis for processing personal data and, where appropriate, an additional condition for processing special category data.

In line with our charter which states that we advance learning and knowledge by teaching and research, the University processes personal data for research purposes under Article 6 (1) (e) of the GDPR:

*Processing is necessary for the performance of a task carried out in the public interest.* Special category data is processed under Article 9 (2) (j):

*Processing is necessary for archiving purposes in the public interest, or scientific and historical research purposes or statistical purposes.*

Research will only be undertaken where ethical approval has been obtained, where there is a clear public interest and where appropriate safeguards have been put in place to protect data.

In line with ethical expectations and in order to comply with common law duty of confidentiality, we will seek your consent to participate where appropriate. This consent will not, however, be our legal basis for processing your data under the GDPR.

##### **How will you use my data?**

Data will be processed for the purposes outlined in this notice.

The data you provide will be used alongside the data of other participants to analyse patterns in the responses. Your data will be stored securely in the University of York, Department of Language and Linguistic Science.

##### **Will you share my data with 3rd parties?**

No. Data will be accessible to the project team at York only. Anonymised data may be reused by the research team for secondary research purposes.

**How will you keep my data secure?**

The University will put in place appropriate technical and organisational measures to protect your personal data and/or special category data. Information will be treated confidentially and shared on a need-to-know basis only. The University is committed to the principle of data protection by design and default and will collect the minimum amount of data necessary for the project. In addition, we will anonymise or pseudonymise data wherever possible.

**Will you transfer my data internationally?**

No. Data will be held within the European Economic Area in full compliance with data protection legislation.

**Will I be identified in any research outputs?**

No real names will be used in any presentations, publications or my dissertation.

**How long will you keep my data?**

Data will be retained in line with legal requirements or where there is a business need. Retention timeframes will be determined in line with the University's Records Retention Schedule.

**What rights do I have in relation to my data?**

Under the GDPR, you have a general right of access to your data, a right to rectification, erasure, restriction, objection or portability. You also have a right to withdrawal. Please note, not all rights apply where data is processed purely for research purposes. For further information see, <https://www.york.ac.uk/records-management/general-dataprotectionregulation/individualsrights/>.

**Questions or concerns**

If you have any questions about this participant information sheet or concerns about how your data is being processed, in the first instance, please contact M.Gabriela Valenzuela, Department of Language and Linguistic Science, University of York, Heslington, York, YO10 5DD, email: [mgvf500@york.ac.uk](mailto:mgvf500@york.ac.uk)

Supervisor's name and details: Peter French, email: [peter.french@york.ac.uk](mailto:peter.french@york.ac.uk)

If you are still dissatisfied, please contact the University's Data Protection Officer at [dataprotection@york.ac.uk](mailto:dataprotection@york.ac.uk)

**Right to complain**

If you are unhappy with the way in which the University has handled your personal data, we ask that you contact us first, to enable us to try to resolve matters. If you remain unhappy, you have a right to complain to the Information Commissioner's Office. For information on reporting a concern to the Information Commissioner's Office, see [www.ico.org.uk/concerns](http://www.ico.org.uk/concerns)

*This study has been reviewed and approved by the Departmental Ethics Committee of the Department of Language and Linguistic Science at the University of York. If you have any questions regarding this, you can contact the chair of the L&LS Ethics Committee, Eytan Zweig, (email: [linguistics-ethics@york.ac.uk](mailto:linguistics-ethics@york.ac.uk); Tel: (01904) 322663).*

## Appendix J

### Perception Test, Consent Form



0% completed

#### 1. Consent

This form is for you to state whether or not you agree to take part in the study. Please read and answer every question. If there is anything you do not understand, or if you want more information, please ask the researcher. [mgvf500@york.ac.uk](mailto:mgvf500@york.ac.uk)

Have you read and understood the information about the study?

[Please choose] ▼

Do you understand that the information you provide will be held in confidence by the researcher, and your name or identifying information about you will not be mentioned in any publication?

[Please choose] ▼

Do you understand that you may withdraw from the study at any time before the end of the data collection session without giving any reason, and that in such a case all your data will be destroyed?

[Please choose] ▼

Do you understand that the information you provide may be kept after the duration of the current project, to be used in future research on language?

[Please choose] ▼

Do you agree to take part in the study?

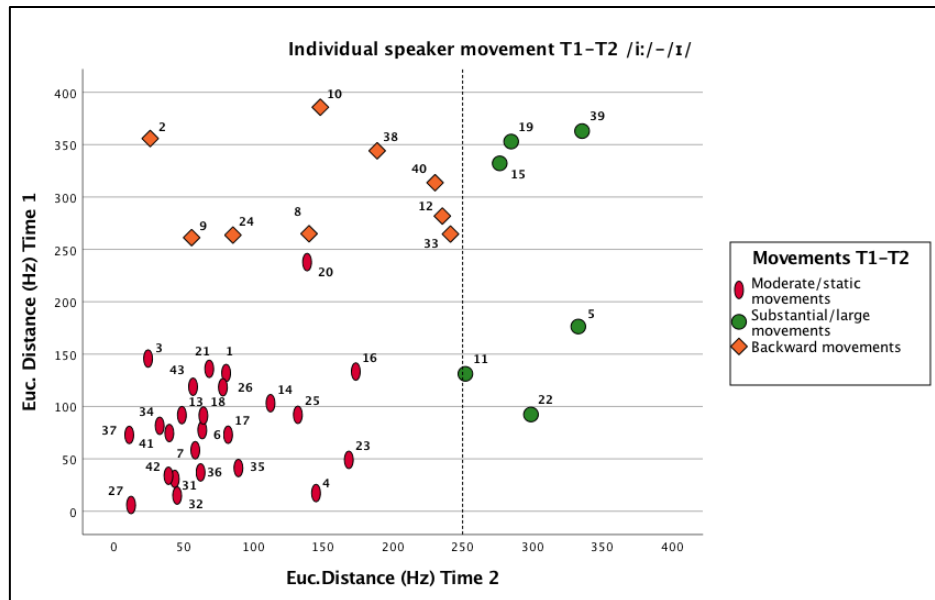
[Please choose] ▼



## Appendix K Individual speakers' trajectories: Connected Speech

**K1 /i:/-/ɪ/**

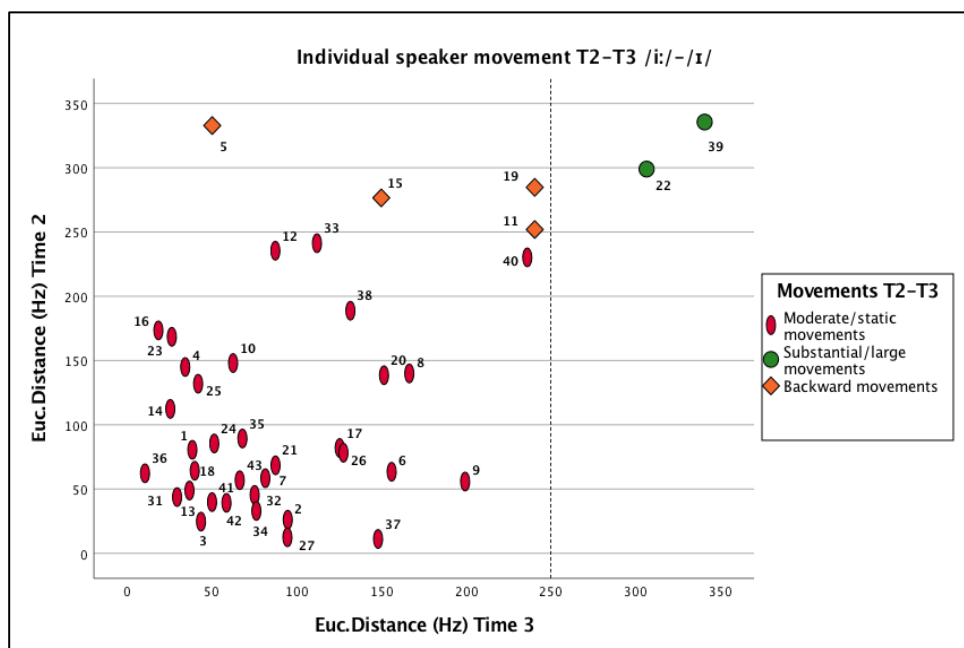
*Vowel distance variations between /i:/-/ɪ/ from Time 1-2*



*Summary Table*

Summary T1-T2	N	%
a) Moderate/static movement (low performers)	25	62.5%
b) Substantial/large movement (high performers)	6	15%
c) Backward movement (low performers)	9	22.5%

Vowel distance variations between /i:/-/ɪ/ from Time 2-3



Summary Table

Summary T2-T3	N	%
a) Moderate/static movement (low performers)	34	85%
b) Substantial/large movement (high performers)	2	5%
c) Backward movement (low performers)	4	10%

**K2 F1-F2**

*Overall results showing the number of speakers producing higher or lower F1 values from Time 1-2*

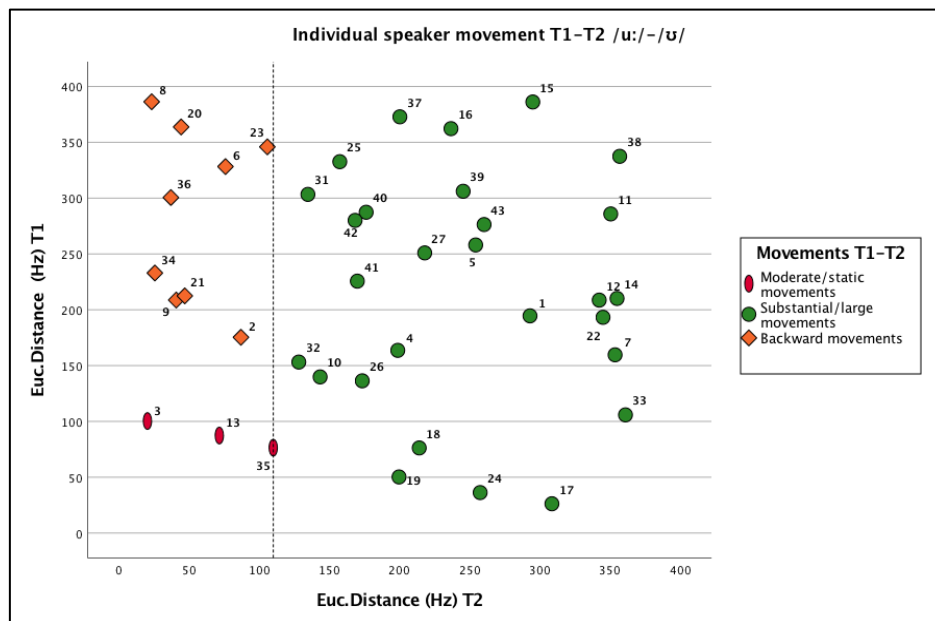
<b>Vowel</b>	<b>N° sp</b>	<b>F1</b>	<b>Vowel</b>	<b>N° sp</b>	<b>F1</b>
/i:/	15	Higher values (avg: 352Hz)	/ɪ /	17	Higher values (avg: 393Hz)
	25	Lower values (avg: 337Hz)		21	Lower values (avg: 378Hz)

*Overall results showing the number of speakers producing higher or lower F2 values from Time 1-2*

<b>Vowel</b>	<b>N° sp</b>	<b>F2</b>	<b>Vowel</b>	<b>N° sp</b>	<b>F2</b>
/i:/	5	Increased values (fronted) (avg: 1866Hz)	/ɪ /	4	Increased values (fronted) (avg: 1712Hz)
	34	Decreased values (central) (avg: 1803Hz)		34	Decreased values (central-back) (avg: 1699 Hz)

### K3 /u:/ and /o/

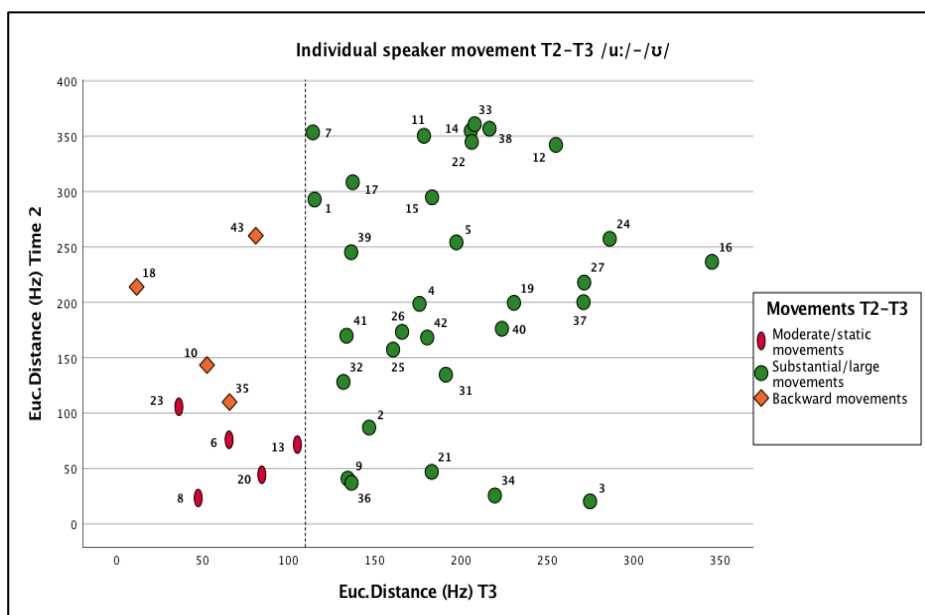
Vowel distance variations between /u:/ and /o/ from Time 1-2



### Summary Table

Summary T1-T2	N	%
a) Moderate/static movement (low performers)	3	7.5%
b) Substantial/large movement (high performers)	28	70%
c) Backward movement (low performers)	9	22.5%

Vowel distance variations between /u:/ and /ʊ/ from Time 2-3



Summary Table

Summary T2-T3	N	%
a) Moderate/static movement (low performers)	4	10%
b) Substantial/large movement (high performers)	32	80%
c) Backward movement (low performers)	4	10%

## K4 F1 and F2

Overall results showing the number of speakers producing higher or lower F1 values from Time 1-2

Vowel	N° sp	F1	Vowel	N° sp	F1
/u:/	22	Higher values (avg: 394 Hz)	/o/	21	Higher values (avg: 426 Hz)
	15	Lower values (avg: 379 Hz)		17	Lower values (avg: 396 Hz)

Overall results showing the number of speakers producing higher or lower F2 values from Time 1-2

Vowel	N° sp	F2	Vowel	N° sp	F2
/u:/	10	Increased (fronted) (avg: 1563 Hz)	/o/	12	Increased (fronted) (avg: 1411 Hz)
	29	Decreased (retracted) (avg: 1371 Hz)		28	Decreased (retracted) (avg: 1271 Hz)

## K5 /ɪ /and /e/

Vowel distance variations between /ɪ /and /e/ from Time 1-2

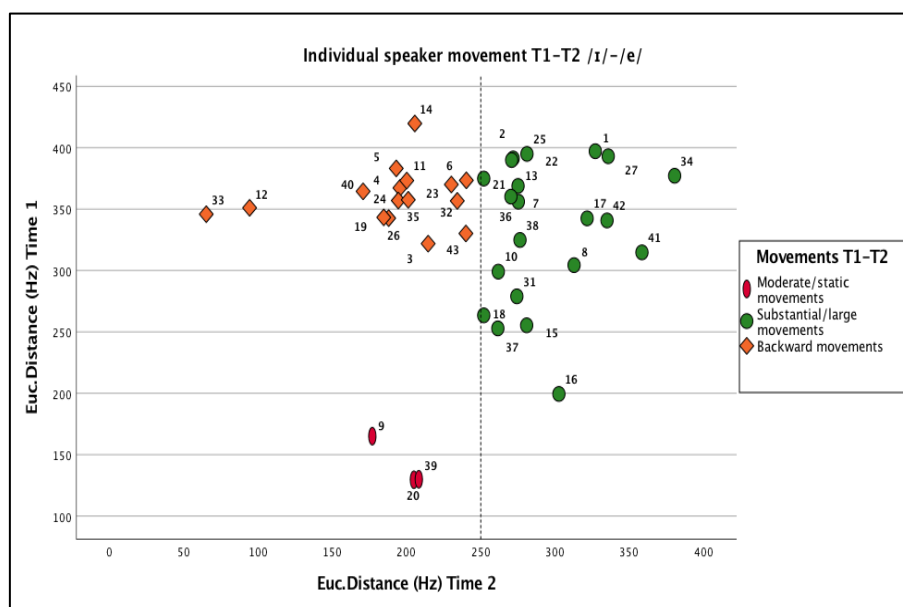


Table Summary

Summary T1-T2	N	%
a) Moderate/static movement (low performers)	3	7.5%
b) Substantial/large movement (high performers)	21	52.5%
c) Backward movement (low performers)	16	40%

Vowel distance variations between /i/ and /e/ from Time 2 - 3

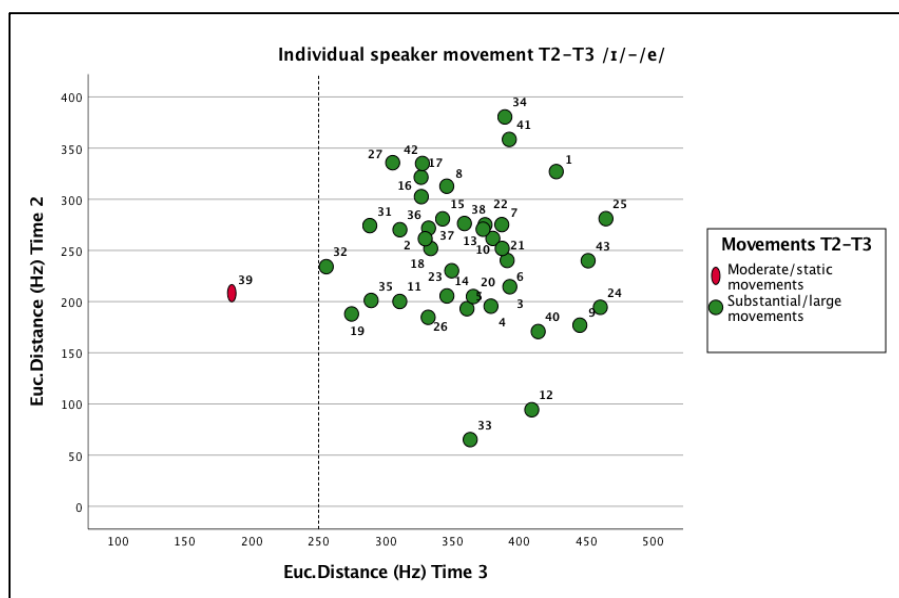


Table Summary

Summary T2-T3	N	%
a) Moderate/static movement (low performers)	1	2.5%
b) Substantial/large movement (high performers)	39	97.5%

**K6 F1- F2**

*Overall results showing the number of speakers producing higher or lower F1 values from Time 1-2*

<b>Vowel</b>	<b>N° sp</b>	<b>F1</b>	<b>Vowel</b>	<b>N° sp</b>	<b>F1</b>
/ɪ/	17	Higher values (avg:393 Hz)	/e/	2	Higher values (avg: 513 Hz)
	21	Lower values (avg: 378 Hz)		38	Lower values (avg: 510 Hz)

*Overall results showing the number of speakers producing higher or lower F2 values from Time 1-2*

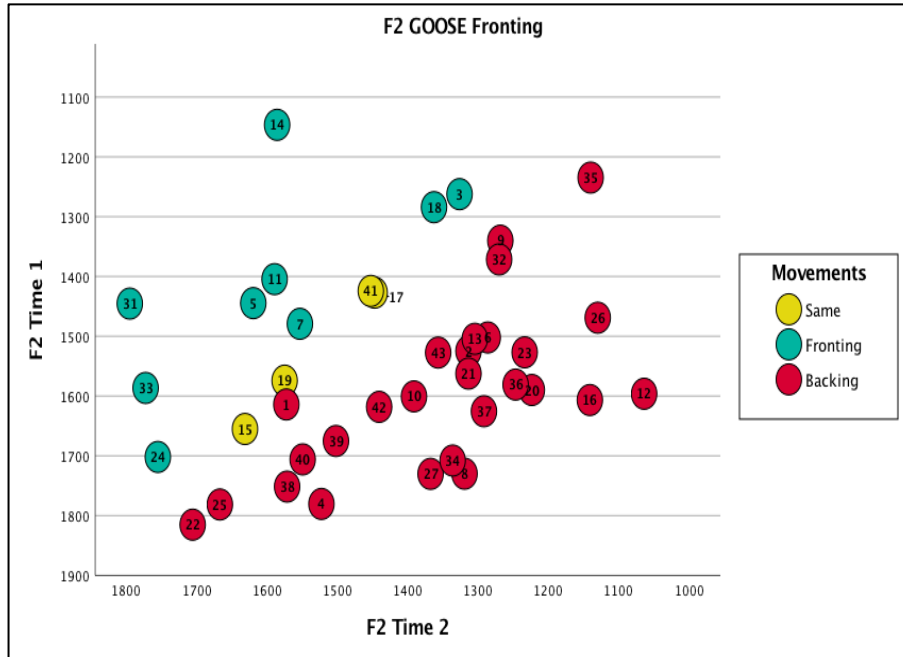
<b>Vowel</b>	<b>N° sp</b>	<b>F2</b>	<b>Vowel</b>	<b>N° sp</b>	<b>F2</b>
/ɪ/	4	Increased values (fronted) (avg: 1712 Hz)	/e/	8	Increased values (fronted) (avg: 1581 Hz)
	34	Decreased values (central-back) (avg: 1699 Hz)		31	Decreased values (central-back) (avg: 1492 Hz)



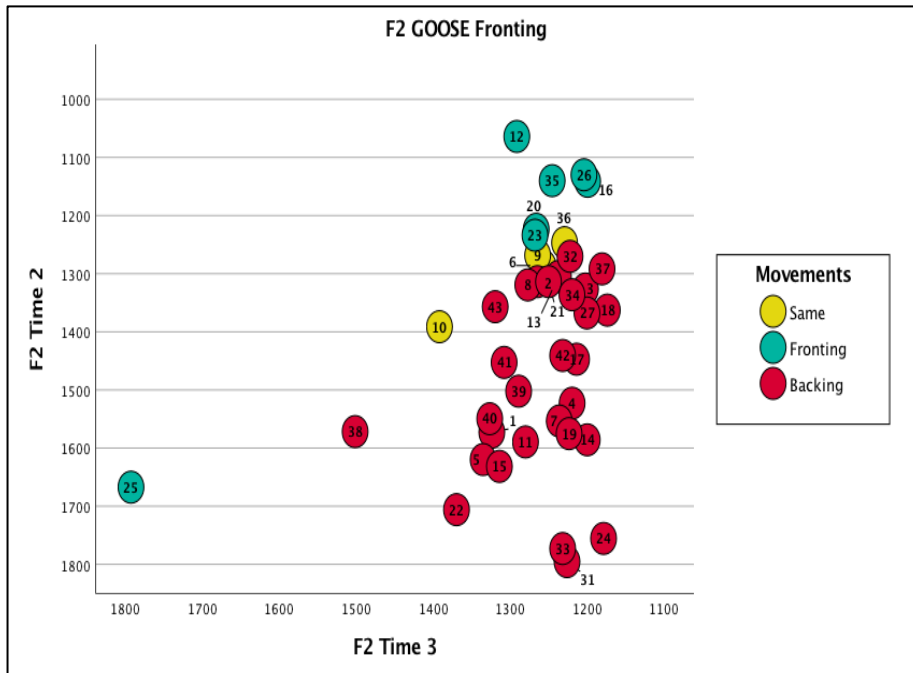
## Appendix L /u: /- /o/ Fronting

### L1 /u:/

Changes in the F2 values for /u:/ between T1 and T2

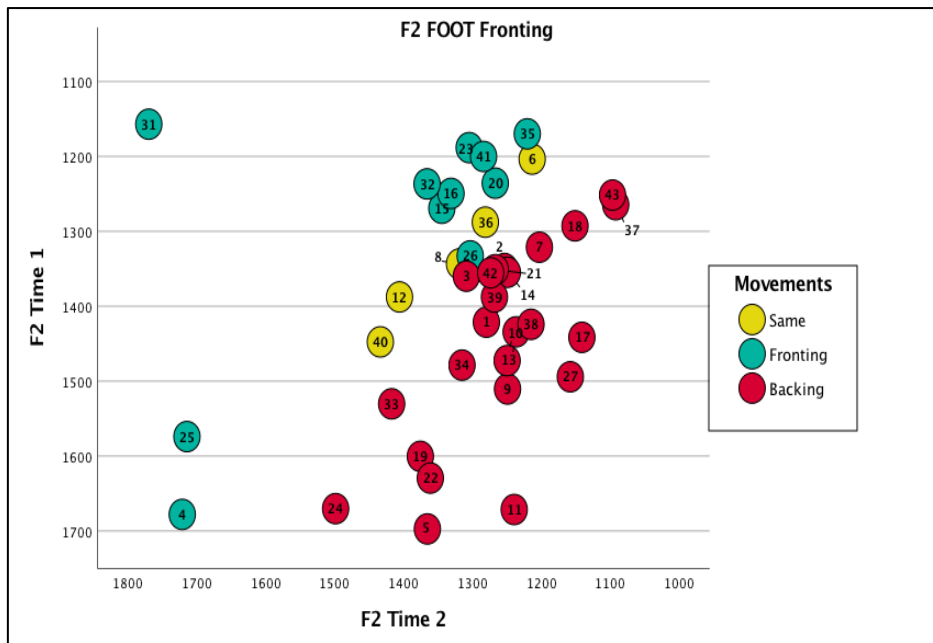


Changes in the F2 values for /u:/ between T2 and T3



## L2 /ʊ/

Changes in the F2 values for /ʊ/ between T1 and T2



Changes in the F2 values for /ʊ/ between T2 and T3

