

Resumptive Pronouns in Baha Arabic: an Experimental Study

Asmaa Alghamdi

Submitted in accordance with the requirements for the degree of
Doctor of Philosophy

The University of Leeds
School of Languages, Cultures, and Societies
Department of Linguistics and Phonetics

November 2020

The candidate confirms that the work submitted is her own and the appropriate credit has been given where reference has been made to the work of others.

This copy has been supplied on the understanding that it is copyright material and that no quotation from this thesis may be published without proper acknowledgement.

The right of Asmaa Alghamdi to be identified as Author of this work has been asserted by her in accordance with the Copyright, Designs, and Patents Act 1988.

© 2011 The University of Leeds and Asmaa Alghamdi

Acknowledgment

I would like to thank the following people, without whom this project would have not been the same. First and foremost, I would like to express my deep gratitude to my primary advisor, Prof. Cecile De Cat, who guided me throughout this project and who always made me feel confident in my abilities. I would like to thank her for her wise guidance, patience, constructive suggestions and support. I would like to express my special thanks to her for giving me her time so generously in all aspects. I have been extremely lucky to have her as my supervisor. I am deeply appreciative of all that she has done for me.

I would also like to thank my supervisor Dr. Diane Nelson who helped me finalize my project. Her suggestions and critiques helped improve the quality of the thesis.

I also wish to thank all of the people whose contributions and availability were crucial in the completion of my research. Special thank to my colleagues at Baha University for their help in recruiting participants for the experimental works included in this thesis. I would also like to extend my thanks to all participants that took part in the study and enabled this research to be possible. I would like to express my special thanks to my sponsor Baha University whose support made this work possible.

I am also deeply thankful to my friend, Afnan Abu Ras, for all the help she provided; her words have been a primary source of support and my stress reliever throughout my PhD journey.

I am heartily thankful to my loving parents who were always by my side despite of being so very far away. This journey would not have been possible without their constant dedication, support and everlasting love. Their words and calm voices always lower my stress during the most difficult times. No words are enough to express my gratitude to them for their everlasting love and support.

I am deeply grateful to my sisters, Fadwa, Munira, Hind, Tagreed, Amal, Sumayya and Sara, and brothers, Saeed and Muhammad, for their constant help and support during this journey. Thank you for always being there for me as friends.

Finally, I would like to thank my loving and caring daughter Mila and my precious son Muhammad for giving me unlimited happiness and pleasure. Their beautiful smiles give me hope and peace, and their nice words always ease my pain. Most sincere appreciation goes to a very special person, my husband Saleh, who stood by me through the good times and bad. He has been extremely supportive of me throughout

this entire process. His enduring love, encouragement, patience and understanding during my journey in pursuing of PhD degree made the completion of this task achievable.

I dedicate this work to my wonderful parents, my husband Saleh, my precious children Mila and Muhammad, and my sisters and brothers.

Abstract

The phenomenon of resumption has been a central topic of debate for both syntacticians and psycholinguists. The debate particularly centres on whether resumption is a syntactic phenomenon or a processing one and whether this characterization differs across languages (and dependencies).

Theoretical and experimental investigations have revealed that resumption exhibits a great deal of variation across (and sometimes within) languages (Sells, 1984; Sells, 1987; Asudeh, 2004; McCloskey, 2006; Asudeh, 2012). This is further complicated by diverging interpretations of the data. Specifically, while the theoretical literature presents resumption as a wide-spread cross-linguistic phenomenon affecting longer distance dependencies and island-violating dependencies, this claim is not confirmed in experimental literature (Farby et al., 2010; Keffala, 2011; Tucker et al., 2019; Perpiñán, 2020, among others).

This dissertation examines the phenomenon of resumption in Baha Arabic, a language with a productive use of resumptive pronouns across different types of dependency structures. It aims to investigate the theoretical literature's argument that resumption is preferred to gaps in certain syntactic configurations (i.e. islands and longer dependencies) and to elaborate on the extent to which resumption in this variety of Arabic differs from resumption in languages like English. Four experimental studies, exploiting both offline and online methods, are conducted. The findings suggest that resumption in Baha Arabic is not a uniform phenomenon, despite the argument that it mainly has a syntactic function; (i) true RPs in *illi*-structures constitute part of the initial derivation in binding dependencies (as morpho-syntactic features of C do not trigger movement) and (ii) intrusive RPs in *wh*-questions are utilized as last resort devices to fix derivation problems when movement is illicit. The availability of intrusive RPs as a syntactic last resort device is restricted to *wh*-questions featuring inherently D-linked fillers 'i.e. *which*-fillers'.

This dissertation, furthermore, examines the extent to which the type of *wh*-filler phrase (*which* vs. *what*) affects the acceptability and processing of island-violating dependencies in Baha Arabic. Although no such claim had previously been made for Arabic varieties, we found that gapped island-violating dependencies with *which*-fillers are accepted more than gapped island-violating dependencies with *what*-fillers. This amelioration effect is interpreted as reflecting an extra-grammatical phenomenon.

Though not the focus of the dissertation, our findings suggest that islands are neither a purely syntactic nor a purely processing phenomenon and that a combination of both cognitive and syntactic constraints contribute to it.

Arabic Transcription Chart

The following is a list of the reading conventions that are used in transcribing the Baha Arabic data.

Consonant Transcription

Arabic script	IPA symbol	Symbol used in this thesis
أ	ʔ	ʔ
ب	b	b
ت	t	t
ث	θ	th
ج	dʒ	j
ح	ħ	ḥ
خ	x	kh
د	d	d
ذ	ð	ð
ر	r	r
ز	z	z
س	s	s
ش	ʃ	š
ص	s ^ʕ	ṣ
ض	d ^ʕ	ḍ
ط	t ^ʕ	ṭ
ظ	ð ^ʕ	ẓ
ع	ʕ	ʕ
غ	ɣ	ġ
ف	f	f
ق	q	q
ك	k	k

ل	l	l
م	m	m
ن	n	n
هـ	h	h
و	w	w
ي	j	y

Vowel Transcriptions

Arabic script	IPA symbol	Symbol used in this thesis
ا	a:	aa
و	u:, o:	uu, oo
ي	i:, e:	ii, ee

Table of Contents

Acknowledgment	iii
Abstract	v
Arabic Transcription Chart	vii
List of Tables.....	xv
List of Figures	xvii
List of Abbreviations.....	xix
Chapter 1 Introduction.....	1
Chapter 2 Background.....	6
2.1 Resumption in the theoretical literature.....	6
2.1.1 Resumption as a homogeneous phenomenon across languages	6
2.1.2 Resumption as a uniform phenomenon across languages	13
2.1.3 Resumption as a homogeneous phenomenon within languages and across constructions	14
2.2 Resumption in the experimental literature.....	17
2.2.1 Experimental literature on resumption in English	17
2.2.2 Experimental literature on resumption in Irish, Hebrew and Arabic	19
2.3 Language under investigation	24
2.3.1 Basic description of Baha Arabic morphosyntax.....	25
2.3.1.1 Word order.....	25
2.3.1.2 Agreement.....	25
2.3.1.3 Case marking	26
2.3.2 Wh-formation and relativisation strategies in Arabic varieties.....	27
2.3.3 Resumption in Highest Subject position.....	30
2.3.4 Pronominal status of resumptive pronouns in Arabic	30
2.4 Conclusion	31
Chapter 3 Resumptive Pronouns in Syntax and Psycholinguistics	33
3.1 Introduction.....	33
3.2 Background on locality constraints	34
3.3 Syntactic accounts	36
3.3.1 FGD formation in syntax.....	36
3.3.2 Islands in syntax.....	37

3.3.3	Resumption in syntax	38
3.3.3.1	Last resort or not last resort	41
3.3.3.2	Complementiser type account.....	43
3.4	WM-based accounts	48
3.4.1	Resource Limitation Theory	48
3.4.2	Resumption in WM-based accounts of locality constraints	51
3.5	Evaluating syntactic vs. WM-based accounts of resumption.....	56
3.5.1	Alexopoulou's (2010) theory: True RPs vs intrusive RPs.....	57
3.5.1.1	Main features of Alexopoulou's (2010) theory	58
3.5.1.1.1	Grammar of true and intrusive RPs is different	58
3.5.1.1.2	Resumption and complexity of processing FGDs.....	58
3.5.1.1.3	Intrusive resumption and islands	60
3.5.1.2	Assumptions on which this theory is based	62
3.5.1.2.1	Grammatical constraints are substantial in theories of island effects	63
3.5.1.2.2	Processing movement and binding dependencies.....	74
3.5.1.2.3	Weak vs strong islands and intrusive RPs.....	76
3.5.1.3	Interim conclusion.....	82
3.6	Type of dependency structure and resumption.....	83
3.7	Summary of our initial assumptions	85
3.8	Research questions	85
3.9	Methodological concerns.....	86
3.9.1	Formal vs informal acceptability judgements.....	86
3.9.2	Gradience in acceptability.....	87
3.10	Conclusion	89
Chapter 4	Acceptability of RPs in Baha Arabic <i>Wh</i>-dependencies.....	90
4.1	Introduction.....	90
4.2	Background.....	90
4.2.1	Resumption and islands.....	90
4.2.2	Resumption and island type.....	93
4.2.3	Resumption and embedding.....	94
4.2.4	Resumption and WM.....	96
4.3	Hypotheses and predictions	96
4.4	Experiment.....	97
4.4.1	Methodology	97
4.4.1.1	Design of the experiment.....	97
4.4.1.2	Materials	101

4.4.1.3	Subjects.....	106
4.4.1.4	Procedures.....	106
4.4.2	Results	107
4.4.2.1	Calculating WM scores.....	107
4.4.2.2	Descriptive data	109
4.4.2.3	Methods of analysis.....	112
4.4.2.4	Statistical analysis.....	113
4.4.2.4.1	Rating score as the dependent variable	113
4.4.2.4.2	Difference as dependent variable.....	124
4.4.3	Discussion	128
4.4.4	Summary of results.....	134
4.4.5	Limitations and questions.....	136
4.4.6	Outstanding questions.....	137
4.5	Conclusion	137
Chapter 5	Processing RPs in Baha Arabic <i>Wh</i>-dependencies	139
5.1	Introduction.....	139
5.1.1	Processing costs of islands.....	140
5.1.2	Processing function of resumption	142
5.1.3	Types of island	145
5.1.4	Cognitive measures of WM capacity and speed of processing.....	146
5.1.5	Summary of research questions and predictions	146
5.1.5.1	Research questions.....	146
5.1.5.2	Predictions.....	146
5.2	Study 2: Processing <i>wh</i>-dependencies with resumption	147
5.2.1	The superadditivity paradigm.....	148
5.2.1.1	Issues in the design of linguistic stimuli.....	149
5.2.2	Revised design.....	152
5.2.3	Participants	156
5.2.4	Procedures	156
5.2.5	Results	158
5.2.5.1	Bare <i>wh</i> -questions.....	158
5.2.5.1.1	Differential RT analysis	159
5.2.5.1.2	Non-differential RT analysis.....	164
5.2.5.2	Relative clauses	169
5.2.5.2.1	Descriptive data	169
5.2.5.2.2	Statistical analysis.....	171

5.2.6	Discussion	173
5.2.7	Limitations.....	174
5.3	Study 3: Processing contextualised D-linked/non-D-linked wh-questions with RPs.....	177
5.3.1	Rationale.....	177
5.3.2	Materials	177
5.3.3	Participants	181
5.3.4	Procedures	181
5.3.5	Analysis and predictions	182
5.3.5.1	Descriptive data	182
5.3.5.2	Statistical analysis	185
5.3.6	Discussion	189
5.3.7	Outstanding questions.....	193
5.4	Conclusion	194
Chapter 6	Discourse Linking and Resumption	195
6.1	Introduction.....	195
6.2	Discourse-linking and resumption	196
6.3	Nature of the D-linking effect	198
6.3.1	D-linking and resumption as syntactic phenomena	201
6.3.2	D-linking and resumption as processing phenomena	204
6.3.3	Differential predictions	207
6.4	Complexity account of <i>which</i>-N phrases.....	208
6.5	Working memory capacity and RPs.....	212
6.6	Study 4: Research questions and hypotheses	213
6.6.1	Methodology	214
6.6.1.1	Subjects.....	214
6.6.1.2	Material	215
6.6.1.3	Procedures.....	219
6.6.2	Results of the acceptability judgment sub-experiment	220
6.6.2.1	Descriptive data	220
6.6.2.2	Statistical analysis	222
6.6.2.2.1	Modelling Procedures	222
6.6.2.3	RT data for the acceptability judgment sub-experiment.....	228
6.6.2.3.1	Descriptive data	228
6.6.2.3.2	Statistical analysis.....	230
6.6.3	RT data in comprehension question data	232

6.6.3.1	Descriptive data	233
6.6.3.2	Statistical Analysis	233
6.7	Discussion.....	237
6.7.1	Syntactic theories.....	238
6.7.2	Processing theories.....	242
6.7.2.1	Resumption as a processing phenomenon.....	242
6.7.2.2	Type of filler phrase	243
6.7.3	Alternative analysis	246
6.7.4	Task Type	248
6.8	Conclusion	249
Chapter 7	General Discussion and Conclusion	251
7.1	Dissertation summary:	252
7.1.1	Summary of the design of the four studies	252
7.1.2	Summary of the main findings	255
7.1.3	Interpretation of results	256
7.1.3.1	Types of RPs in BA	256
7.1.3.1.1	RPs in <i>illi</i> -structures.....	256
7.1.3.1.2	RPs in <i>wh</i> -questions.....	256
7.1.3.2	Interaction between type of <i>wh</i> -filler and islands.....	261
7.1.4	The effect of the behavioural tasks	261
7.1.5	Conclusion.....	263
7.2	Consequences for processing accounts of FGDs	264
7.2.1	Summary of existent processing accounts of FGDs	264
7.2.2	Implication of findings on accunts of processing FGDs.....	266
7.3	Consequences for accounts of island phenomena.....	268
7.4	Variation across Arabic varieties.....	272
7.5	Future work	274
References.....		275
APPENDIX A	Experimental Sentences	290
	Sentences included in Study 1.....	290
	Sentences included in Study 2.....	306
	Sentences included in Study 3.....	325
	Sentences included in Study 4.....	339
APPENDIX B	Language Background Questionnaire (in English)	346

APPENDIX C Language Background Questionnaire (in Arabic)	348
APPENDIX D Information Sheet (English Version)	350
APPENDIX E Information Sheet (Arabic Version)	352
APPENDIX F Consent Form (English Version)	354
APPENDIX G Consent Form (Arabic Version)	356

List of Tables

Table 4-1 Conditions involved in the first experiment.	101
Table 4-2 Results of the backward digit span task of one subject.....	107
Table 4-3 Comparison between mean rejection of core experimental items with fillers.	111
Table 4-4 Summary of the optimal Ordinal Cumulative Link Mixed Model (clmm) model of the Rejection ratings.	114
Table 4-5 Coefficients of random effects.....	116
Table 4-6 Threshold coefficients.....	117
Table 4-7 Means of difference scores calculated across conditions as (acceptability score of Gap - acceptability score of RP).....	125
Table 4-8 Summary of the parametric terms from the Generalised Additive Model fitted to Difference data.....	126
Table 4-9 Coefficients of random factors	127
Table 5-1 Summary of the Generalised Additive Model fitted to Difference data. Dependent variable: Difference: RT(gap) – RT (RP).....	162
Table 5-2 The mean raw RTs at integration site and the mean average raw RTs at spillover region across conditions.....	165
Table 5-3 Summary of the non-parametric additive quantile regression model of RT data: quantile '0.25'.....	168
Table 5-4 The mean raw RTs at integration site and the mean average raw RTs at spillover region across conditions.....	170
Table 5-5 Summary of the Generalised Additive Model fitted to Difference data...	172
Table 5-6 Summary of the Generalised Additive Model fitted to Difference data. Dependent variable: Difference: RT(gap) – RT (RP).....	186
Table 6-1 Comparison between mean acceptability of core experimental items with fillers.	221
Table 6-2 Summary of the optimal Ordinal Cumulative Link Mixed Model (clmm) of the Acceptability ratings	223
Table 6-3 Coefficients of random effects.....	223

Table 6-4 Threshold coefficients.....	224
Table 6-5 The mean raw RTs at integration site and the mean average raw RTs at spillover region across conditions.....	229
Table 6-6 Summary of the optimal Generalized Additive Model (GAM) of the RT data in the Acceptability Judgment sub-experiment.....	231
Table 6-7 The mean raw RTs at integration site and the mean average raw RTs at spillover region across conditions.....	233
Table 6-8 Results of Nested Model comparisons as revealed by compareML method.	235
Table 7-1 Comparative summary of the design of the four experimental studies. .	253

List of Figures

Figure 4-1 Rejection ratings showing mean and 95% confidence intervals across conditions and dependency structures, with and without RP.	109
Figure 4-2: Mean rejection ratings with confidence intervals for the different conditions, each with and without RP, across the three dependency structures.	111
Figure 4-3 Results of the Ordinal Cumulative Link model showing the probability curves for each of the seven ratings in each condition (represented with vertical lines), along with Confidence Intervals.	118
Figure 4-4 Variability within participants' rating of bare <i>wh</i> -questions with RPs in the NonCrossed_Island and Crossed-NonIsland conditions.	120
Figure 4-5 Inter-individual variations in the rating of bare <i>wh</i> -questions with RPs in the NonCrossed_Island and Crossed-NonIsland conditions.	121
Figure 4-6 Inter-individual variations in the rating of bare <i>wh</i> -questions with RPs in the NonCrossed_Island and Crossed-NonIsland conditions.	122
Figure 4-7. Predicted values of lexicalisation across the different conditions in the acceptability judgment study.	124
Figure 4-8 Difference scores calculated as (gap – RP) in <i>illi</i> -structures and BareWH.	128
Figure 5-1 Means with confidence intervals of differential RTs at the spillover region across different conditions in bare <i>wh</i> -questions.	160
Figure 5-2 Distribution of differential RTs.	161
Figure 5-3 The summed effects of the parametric terms in the GAM model.	163
Figure 5-4 Means with confidence intervals of raw RT scores in gap-bare <i>wh</i> -questions and RP-bare <i>wh</i> -questions across conditions.	165
Figure 5-5 Distribution of box-cox non-differential RT data.	167
Figure 5-6 Predicted values of RT scores for Tail in bare <i>wh</i> -questions. Reference level: gap.	168
Figure 5-7 Means with confidence intervals of raw RT scores in gap relatives and RP relatives.	170

Figure 5-8 Distribution of box-cox transformed RTs at the spillover region.....	171
Figure 5-9 Distribution of differential RTs before removing points with an absolute value greater than 2000 ms.	183
Figure 5-10 Distribution of differential RTs after removing points with an absolute value greater than 2000 ms.	184
Figure 5-11: Means with confidence intervals of differential RTs across different conditions in D-linked and non-D-linked questions.	185
Figure 5-12 The summed effects of the parametric terms in the GAM model.....	187
Figure 6-1: Acceptability ratings showing mean and 95% confidence intervals across structures with different fillers (What/Which), and salience (Salient vs non-Salient), with and without RP.....	221
Figure 6-2: Results of the Ordinal Cumulative Link model showing the probability of each of the seven ratings in each condition (represented with vertical lines), along with Confidence Intervals (shaded).	225
Figure 6-3: Variability in the ratings received by each lexicalisation in <i>which</i> -questions with gaps.	226
Figure 6-4 Variability in the rating of <i>which</i> -questions with gaps by subjects.....	227
Figure 6-5 Variability in the rating of <i>which</i> -questions with gaps by items across subjects.	228
Figure 6-6 Average of raw RT of the three words in the spillover region, showing mean and 95% confidence intervals across questions with different types of fillers (What vs Which) with and without RP.	229
Figure 6-7 Distribution of box-cox-transformed RT data.....	231
Figure 6-8 Distribution of box-cox RT data.....	234
Figure 6-9 Plot with the average raw RT at the three-word spillover region, showing mean and 95% confidence intervals across questions with different types of fillers (What vs Which) with and without RP, in the two different tasks.	236

List of Abbreviations

The following is a list of the abbreviations used in the thesis:

- 1. First person
- 2. Second person
- 3. Third person
- F. Feminine
- M. Masculine
- S. Singular
- PL. Plural
- DU. Dual
- NOM. Nominative
- ACC. Accusative
- GEN. Genitive
- SPEC. Specifier
- COMP Complementiser
- WM Working Memory
- RLT Resource Limitation Theory
- BA Baha Arabic
- RC. Relative clause
- OP. Operator
- RP Resumptive Pronoun
- FGD Filler Gap Dependency
- WM Working Memory
- PS Processing Speed
- EPP Extended Projection Principle
- vP Verb phrase
- SVO Subject-Verb-Object

LF Logical Form
IP Inflectional Phrase
DP Determiner Phrase
RT Reaction Time
PASS Passive

Chapter 1 Introduction

A major property of human languages is the existence of long-distance dependencies. These structures contain two related elements, one of which does not appear in its canonical position, and depends on the other for its grammatical features and interpretation. The two elements can be separated by one or more clause boundaries. For example, a dependency is established in (1) between the *wh*-word *what*, referred to as a ‘filler’ in the psycholinguistic literature, and the empty position of the verb’s argument, indicated by the underscore and referred to as a ‘gap’. Hence, the filler phrase inherits its theta-role from the gap, where it is assigned by the predicate *write*; as a result, the filler phrase is interpreted as the verb’s semantic argument.

1. **What_i** does Jane write ____i ?

What_i does John think that Jane wrote ____i ?

What_i does John believe that Sara thinks that Jane wrote ____i ?

Many different types of long-distance dependencies exist in addition to *wh*- questions, some of which are illustrated in (2).

2. I met **the girl_k** that you admire ____k (relativisation)

That car_k, I want to drive ____k (topicalisation)

This is **the dress_k** that I want to buy ____k (clefting)

All of the above dependency structures share the property of having a filler phrase that does not appear in its canonical position but depends on the gap for its grammatical features and interpretation, regardless of how many clauses separate them. These structures are referred to as filler-gap dependencies (FGDs).

Some languages make use of a different strategy for building long distance dependencies in which a filler phrase binds a pronominal element instead of a gap at dependency tail, as in (3). These pronouns are labeled as resumptive pronouns (RPs).

3. a. I-bint ?illi šufti=ha
the-girl that saw=her
‘the girl that you saw’

(Palestinian Arabic; Shlonsky, 1992, pp.444–445)

- b. ha-ʔiš še- xašavti al=av
 the-man that-1 thought about=him
 'the man that I thought about'

(Hebrew; Shlonsky, 1992, pp.444–445)

In languages like English, on the other hand, the use of the resumptive strategy to build long distance dependencies is described as marginal, and is restricted to cases where gaps are either not allowed or dispreferred due to violation of grammatical constraints or due to an increase in the processing load on Working Memory (WM) resources (Alexopoulou, 2010; McCloskey, 2017). See (4) and (5).

4. a. *These are the things that we do not know what __ are.
 b. ?These are the things that we do not know what they are.

(Chacón, 2015, p.92)

5. There is this new kind of floor that one of the studios in New York that I danced in has it.

(Blythe, 2016, p.159, taken from Kroch, 1981)

The phenomenon of resumption has been a central topic of debate for both syntacticians and psycholinguists, as it exhibits a great deal of variation across and within languages. The debate particularly centers on whether resumption is a syntactic phenomenon or a processing one and whether this characterization differs across languages (and dependencies). Three alternative possibilities have been proposed in the literature on resumption. One is that resumption is a homogenous, narrow syntax phenomenon in certain languages (possibly restricted to particular dependencies) (i.e., grammatical RPs) while it falls at the interface of syntax and processing (i.e., intrusive RPs) in others (Alexopoulou, 2010; McCloskey, 2017). Alternatively, resumption might be a uniform (syntactic or processing) phenomenon cross-linguistically, operating as a last resort device (i.e., intrusive RPs) (Shlonsky, 1992; Ariel, 1999; Hawkins, 2004). Finally, it is possible that RPs have different functions within the same language, depending on the structure where they occur (non-island vs. island) (Erteschik-Shir, 1992; Sterian, 2016).

The debate is further complicated by divergence of findings between the theoretical and experimental literature on resumption. Experimental investigations on resumption in many languages, including Hebrew (Farby et al., 2010), Modern Standard Arabic (Tucker et al., 2019), Spanish (Perpiñán, 2020) and English (Keffala, 2011, among others), are not fully consistent with the claims of the theoretical literature. Specifically, while theoretical literature claims that resumption is common across languages in certain syntactic configurations, termed ‘islands’, this claim is not confirmed in experimental literature.

Furthermore, corpus-based studies in languages where resumption is described as grammatical have revealed that resumption is largely dispreferred in situations where processing demands are not high (Ariel, 1999; McCloskey, 2017). This is not consistent with the theoretical assumption that gap-dependencies and RP-dependencies are equally acceptable in these languages when there is no violation of island constraints.

In this dissertation, we aim to contribute to the cross-linguistic literature on resumption through experimental investigations of resumption in Baha Arabic. Similarly to other varieties of Arabic, Baha Arabic is described as a language with a productive use of resumption across different types of dependency structures. We tested several types of dependency structures (i) to further investigate the argument in the theoretical literature that resumption saves island effects and (ii) to elaborate on the extent to which resumption in this variety of Arabic differs from resumption in languages like English.

Four experimental studies, exploiting both online and offline methods, were conducted for these purposes. We detected a significant improvement in the acceptability of island-violating dependencies when an RP appears at dependency tail instead of a gap in three dependency structures: relative clauses and cleft *wh*-questions (both feature the complementizer *illi*) and *which*-questions. Island-violating dependencies in *what*-questions, on the other hand, are not affected by whether the dependency terminates in a gap or in a pronoun.

The high acceptance rate of RPs in island-violating dependencies both in *illi*-structures and *which*-questions suggests that the impact of resumption is at the grammatical level, and that RPs in these structures are interpreted as bound variables as traces.

Furthermore, we observed that island-violating *which*-questions and island-violating *illi*-structures with gaps and RPs behave differently in several aspects (both in terms of processing and acceptability). These observations lead us to conclude that resumption is not a uniform phenomenon in BA, despite the fact that it mainly has a syntactic function:

- (i) True RPs in *illi*-structures (relative clauses and cleft *wh*-questions) constitute part of the initial derivation in binding dependencies (as morpho-syntactic features of C do not trigger movement) (McCloskey, 2002; Alexopoulou, 2010).
- (ii) Intrusive RPs in D-linked *which*-questions are utilized as last resort devices to fix syntactic derivation problems (Shlonsky, 1992; Aoun, 2000).

Another issue that this dissertation addresses is the extent to which the type of *wh*-filler affects the acceptability and processing of FGDs that involve the violation of an island constraint. For instance, it has been claimed that the violation of island constraints is alleviated when the filler phrase is D-linked (or semantically and structurally complex) (i.e. *which*-fillers), as in (6b), in contrast to when it is non-D-linked (or semantically and structurally simple) (i.e. *what*-fillers), as in (6a).

6. a. *What do you believe the claim that Mary saw ___ ?
b. Which movie do you believe the claim that Mary saw ___ ?

(Goodall, 2017, p.66)

This hypothesis has been tested in several languages, including English (Hofmeister and Sag, 2010; Alexopoulou and Keller, 2013; Goodall, 2014), Greek (Alexopoulou and Keller, 2013) and German (Freitag et al., 2013; Freitag and Repp, 2015). Arabic syntacticians make no such argument for Arabic varieties: the commonplace view is that *which*-questions with violations of island constraints are not acceptable unless a resumptive pronoun appears at dependency tail (Aoun et al., 2009). In this thesis, we tested this hypothesis in BA and detected a significant and reliable impact of type of filler on acceptability and processing of island-violating dependencies with gaps. In line with Hofmeister and Sag (2010), Goodall (2014) and Goodall (2017), we argued that this effect represents an extra-grammatical effect.

The phenomena of resumption and D-linking (or semantic and structural complexity of filler phrases) are relevant to the discussion about the source of island effects. We argue that the findings of this dissertation highlight the importance of both syntactic and processing-related factors in the accounts of the source of island effects.

Empirically, this dissertation contributes both to the documentation of an understudied variety of Arabic, namely Baha Arabic, and to the cross-linguistic literature on RPs. It further contributes to the debate regarding the source of island effects.

The structure of the dissertation is as follows: Chapter 2 presents an overview of resumption typology, as described in the theoretical and experimental literature. Chapter 3 reviews the syntactic and psycholinguistic accounts of the role of RPs in syntax and the processing of long-distance filler-gap dependencies. Chapter 4 presents an acceptability judgment of resumption in different types of wh-dependencies in BA, and Chapter 5 presents two online studies for the processing of resumption in different types of wh-dependencies in BA. Chapter 6 presents an online study that investigates the nature of the interaction between discourse properties of filler phrases and resumption in BA island-violating wh-questions. Chapter 7 presents a general discussion for the results of the four experimental studies and concludes the dissertation.

Chapter 2 Background

Filler resumptive dependencies have been found in several languages. However, theoretical and experimental investigations have revealed that resumption exhibits a great deal of variation across (and sometimes within) languages (Sells, 1984; Sells, 1987; Asudeh, 2004; McCloskey, 2006; Asudeh, 2012). The dimensions and nature of this variation remain controversial. This chapter will provide a general brief description of the variation in the behaviour of resumptive pronouns across and within languages, as described in the theoretical and experimental literature.

The remainder of this chapter is structured as follows. First, we will introduce different perspectives on the nature and scope of the variation in the phenomenon of resumption in theoretical literature. We will next review the experimental literature on resumption, after which we will present the language under investigation and explain why this is an interesting language to study in relation to resumption. The chapter ends with a brief description of the morpho-syntactic characteristics of the language under investigation.

2.1 Resumption in the theoretical literature

The study of resumption in the theoretical literature aims to identify the extent to which the structural properties of RP dependencies differ from those of gap dependencies. In this section, we will introduce different views on the nature of variation in the phenomenon of resumption across languages. We will first introduce the common view, which assumes that there are two distinct types of languages with two distinct types of RPs. Next, we will introduce the perspectives that treat RPs cross-linguistically as a universal phenomenon. Finally, we will present the view that assumes that distinct types of RPs are used across different constructions within the same language.

2.1.1 Resumption as a homogeneous phenomenon across languages

A clear-cut distinction has been drawn between two types of languages depending on how resumption is employed. According to McCloskey (2006), in one set of languages the resumptive strategy is grammatically licensed as a strategy for building syntactic dependencies in line with the gap strategy; i.e. the two mechanisms for establishing long-distance dependencies (the resumptive and the gap strategies) coexist in these languages. Hebrew, Irish and varieties of Arabic are representative of this class of

languages. The example below illustrates that RPs and gaps alternate freely in Irish (1), Hebrew (2) and Iraqi Arabic (3); in these examples, the syntactic dependency is resolved either with a gap or with an RP.

1. a. an bheirt a bhí siad ag iarraidh
the two COMP-FG be.PST they PROG try
a shábháil
save.N-FIN

‘The two that they were trying to save’

b. an bheirt a raibh siad ag iarraidh
the two COMP-RP be.PST they PROG try
iad a shábháil
them save.N-FIN

‘The two that they were trying to save them’

(Irish; McCloskey, 2017)

2. a. ha- ʔiš še- raʔiti ____
The-man that-I saw ____

‘The man that I saw’

b. ha-ʔiš še- raʔiti ʔoto
The-man that-I saw him

‘The man that I saw him’

(Hebrew; Shlonsky, 1992, p.444)

3. a. Iman ya: riɖɖʒa:l ša:fet ____ bi-l-ħafla
Iman which man saw.3FS ____ at-the-party

‘Which man did Iman see ____ at the party?’

b. Iman ya: riɖɖʒa:l ʃa:fet=**eh** bi-l-ħafla
Iman which man saw.3FS=**3MS** at-the-party

‘Which man did Iman see [**him**] at the party?’

(Iraqi Arabic; Sterian, 2016, p.109)

However, in some syntactic structures, gaps are not allowed to occur at the tail of a dependency structure. These constructions are known as ‘islands’ (see Chapter 3 for a detailed discussion of the phenomenon of island effects). In brief, when a dependency resolves inside island structures, gaps are ungrammatical. In this case, resumptive pronouns obligatorily occur to resolve the dependency structure. The following examples from Hebrew (4), Iraqi Arabic (5), and Irish (6) all include violations of island constraints; here, gap dependencies are not acceptable, while RP dependencies are acceptable.

4. a. With RP:

raʔiti ʔet ha-yeled še Dalya makira ʔet ha-ʔiša
saw.1S et the-boy that Dalya knows et the-woman
še ʔohevet oto
that loves him

‘I saw the boy that Dalya knows the woman who loves him.’

b. With Gap:

*raʔiti ʔet ha-yeled še Dalya makira ʔet ha-ʔiša še
saw.1S et the-boy that Dalya knows et the-woman that
ʔohevet____
loves____

‘I saw the boy that Dalya knows the woman who loves.’

(Hebrew; Borer, 1984)

5. a. With RP:

ya: Su:ra.F li-bni=ha Samer ysʔal ida
which picture.F of-son=her Samer ask.3MS if
kull mraya šagagat=**ha**
every woman tore.3FS=**3F**

‘Which photo of her son did Samer wonder if every woman tore [it]?’

b. With Gap:

*ya: Su:ra.F li-bni=ha Samer ysʔal ida

which picture.F of-son=her Samer ask.3MS if
kull mraya šagagat_____

every woman tore.3FS_____ '

'Which photo of her son did Samer wonder if every woman tore ___?'

(Iraqi Arabic; Sterian, 2016)

6. a. With RP:

Sin teanga aN mbeadh meas agam ar
that a.language COMP would be respect at me on

duine ar bith aL tá ábalta i a

person any COMP is able it to

labhairt

speak

'That's a language that I would respect anyone who could speak it.'

b. With Gap:

*an fear aL phóg mé an bhean aL phós

the man COMP kissed I the woman COMP married

'the man who I kissed the woman who married'

(Irish; McCloskey, 1979).

In essence, theoretical syntactic accounts assume that RP dependencies and gap dependencies have different underlying syntactic structures: i.e. the derivational properties of both types of dependencies in syntax are different. It is commonly assumed that the gap strategy is derived by syntactic successive-cyclic movement, while the RP strategy is derived by syntactic binding relations; the former is known to be sensitive to island effects, while the latter is not (McCloskey, 2006; Alexopoulou, 2010; Asudeh, 2012). See section 3.3 in Chapter 3 for a more detailed illustration.

In most varieties of Arabic, the resumptive strategy is the only available option for creating syntactic dependencies in certain dependency structures, regardless of whether a crossed island structure is present; the below examples from Iraqi Arabic (7), Jordanian Arabic (8) and Palestinian Arabic (9) illustrate this point. The lack of

resumptive pronouns in these examples renders these sentences unacceptable, despite the fact that these dependency structures do not involve gaps inside islands.

7. a. With Gap:

*il-ka:teb illyi šeft ____ eb-be:t Suha tša:n
the-writer whom saw.1S____ at-house Suha was
Khalil Jubran
Khalil Jubran

'The writer whom I saw ____ in Suha's house was Khalil Jubran.'

b. With RP:

Il-ka:teb illyi šeft=**ah** eb-be:t Suha tša:n
the-writer whom saw.1S=**him** at-house Suha was
Khalil Jubran
Khalil Jubran

'The writer whom I saw [**him**] in Suha's house was Khalil Jubran.'

(Iraqi Arabic; Sterian, 2016)

8. a. With Gap:

*I-ktab illi štarayt ____ mbariḥ ḍaḥa
the-book.M which bought.1S____ yesterday lost

'The book that I bought ____ yesterday is lost.'

b. With RP:

I-ktab illi štarayt-uh mbari ḥ ḍaḥa
the-book.M which bought.1S=it yesterday lost

'The book that I bought [it] yesterday is lost.'

(Jordanian Arabic; Guilliot and Malkawi, 2009)

9. a. With Gap:

*I-bint ijilli (you-f) šufti-____
the-girl that saw-gap

b. With RP:

l-bint ijilli šufti=ha
the-girl that (you-f) saw=her
'the girl that you saw'

(Palestinian Arabic; Shlonsky, 1992, p.445)

However, some dependency structures in this set of languages do not allow resumptive pronouns to appear at the dependency tail. For instance, Aoun et al. (2009) argued that, in Lebanese Arabic, resumptive pronouns are not allowed to occur with *wh*-questions when the filler phrase is 'what', as in (10). Similarly, Sells argued that resumptive pronouns are not acceptable in simple *wh*-questions in Hebrew, as in (11).

10. *šu štarayt-i
what bought.2SF=it
'What did you buy it?'

(Lebanese Arabic; Abdel Razaq, 2011)

11. *mi nifgaSta ito
who you-met with=him
Who did you meet with him?

(Hebrew; Sharvit, 1999)

Moving to the other set of languages, it is claimed that the only strategy licensed by the grammar for establishing syntactic dependencies is the gap strategy; i.e. that the resumptive strategy does not constitute a part of the grammatical system in these languages. In other words, the syntactic operation involved in deriving RP dependencies in languages like Arabic, Irish and Hebrew (i.e. the pronoun operator-binding relation) is not licensed in the syntax of this set of languages. English is representative of this set of languages. For example, the use of a RP in a structure like (12) is considered unacceptable.

12. the girl that I met (*her) at the party

(Alexopoulou, 2010)

Hence, it is assumed that the syntactic system of this set of languages does not support the creation of syntactic dependencies with resumptive pronouns and that the only option available in the grammar for creating syntactic dependencies is to use

gaps. Therefore, resumption is argued not to be syntactically licensed in English. Despite this, resumption in English has been attested in corpus studies (Kroch, 1981; Prince, 1990; Blythe, 2016), as illustrated in the following examples:

13. Let's get to our first guest, who I asked and was so delighted that **he** could make it.

(Prince, 1990, p.482, taken from Kroch, 1981)

14. That asshole X, who I loathe and despise the ground **he** walks on, pointed out that...

(Prince, 1990, p.483, taken from Kroch, 1981)

15. There is this new kind of floor that one of the studios in New York that I danced in has **it**.

(Blythe, 2016, p.159, taken from Kroch, 1981)

To account for the puzzling fact that these expressions continue to be used despite their ungrammaticality, it is generally argued that resumption in English is used as a processing (repair) device (Erteschik-Shir, 1992; Asudeh, 2004; Asudeh, 2012; Hofmeister and Norcliffe, 2013; Ackerman et al., 2018, among others).

Advocates of this view suggest that the use of resumption in English is restricted to cases in which processing long-distance dependencies is too demanding; in particular, it is proposed that use of RPs is limited to cases in which resolving FGDs is difficult or impossible; for instance, due to an increase in processing load, such as when a dependency involves multiple levels of embedding. See e.g. (16) below (Kroch, 1981; Prince, 1990; Erteschik-Shir, 1992; Asudeh, 2004; Asudeh, 2012).

16. I just saw that girl who Long John's claim that **she** was a Venusian made all the headlines.

(Ross, 1986, p.260)

Furthermore, use of resumption in English is common inside island structures, where gaps are not allowed (Ross, 1967; Erteschik-Shir, 1992; Asudeh, 2012); see (17) and (18) below.

17. a. Didn't that guy who the Game Warden and **him** had seen a flying saucer crack up?

b. Palmer is a guy who for **him** to stay in school would be stupid.

(Ross, 1967, p.432)

18. This is the girl that Peter said that John thinks that yesterday his mother had given some cakes to **her**.

(Asudeh, 2004)

However, it has also been reported that RPs inside island structures in English may improve, but not fully repair, island-violating dependencies (Alexopoulou, 2010). Hence, it is claimed that RPs can facilitate the processing of FGDs, making it possible to process and interpret these complex structures regardless of their grammatical status. See (19).

19. a. *These are the things that we do not know what ___ are.
b. ?These are the things that we do not know what they are.

(Chacón, 2015, p.92)

Chao and Sells (1983) refer to resumptive pronouns in the former set of languages as 'true RPs', while those in the second set of languages are referred to as intrusive pronouns. In the theoretical literature, the dominant view assumes that resumption in these two sets of languages instantiate two distinct phenomena: in the former set of languages, true resumption is considered to be a strategy for creating long-distance dependencies in syntax; in the latter set of languages, intrusive resumption is viewed as an extra-grammatical 'repair' mechanism employed to overcome difficulty in establishing syntactic dependencies (Erteschik-Shir, 1992; Asudeh, 2004; Alexopoulou and Keller, 2007; Alexopoulou, 2010; Polinsky et al., 2013).

2.1.2 Resumption as a uniform phenomenon across languages

Shlonsky (1992) argued that all resumptive pronouns in both intrusive resumption languages (such as English) and true resumption languages (such as Hebrew and Palestinian Arabic) have the same representational properties. In other words, Shlonsky (1992) assumed that RPs cross-linguistically are intrusive pronouns that are used as last resort expressions in contexts where syntax blocks movement dependencies. Such contexts include island structures, which represent a universal phenomenon that blocks syntactic movement. Shlonsky's (1992) last resort account accordingly claims that resumption inside islands is cross-linguistically common.

To account for why RPs occur in non-island contexts in true resumption languages, Shlonsky (1992) argued that a particular type of relative complementiser, such as *illi*

in Palestinian Arabic, blocks movement operations. In such cases, a resumptive pronoun must appear as a last resort strategy to resolve the dependency. Languages such as English, on the other hand, lack such complementisers. Specifically, Shlonsky's (1992) account assumes that true resumption languages have more contexts that do not allow gaps to appear inside them when compared to intrusive resumption languages. This account accordingly assumes that the use of RPs in English is restricted to island-violating dependencies, while in true resumption languages, RPs can appear in non-island contexts with certain types of complementisers.

Ariel (1999) and Hawkins (2004) proposed processing-based accounts for resumptive pronouns, regardless of whether they occur in a true or an intrusive resumption language. In fact, these authors acknowledged the differences in behaviour between RPs in true and intrusive resumption languages in terms of their grammatical status. However, they did not propose that these two types of RPs have different representational properties; instead, they claimed that true resumption arises through the conventionalisation of intrusive resumptive pronouns into the grammar. In summary, these accounts argued that, across languages, RPs are preferred to gaps when the complexity of processing FGDs is increased.

2.1.3 Resumption as a homogeneous phenomenon within languages and across constructions

A different perspective assumes that resumptive pronouns inside islands are of the intrusive type across languages (Erteschik-Shir, 1992; Rouveret, 2011; Sterian, 2016). However, languages differ in terms of the grammatical status of resumptive pronouns in non-island, easy-to-process contexts. This view implies that a distinction is made in true resumption languages between resumptive pronouns in island and non-island contexts.

Erteschik-Shir (1992) argued that Hebrew features two types of RPs, depending on the structure in which they occur. Following Doron (1982), Erteschik-Shir (1992) argued that resumptive pronouns used in non-island, simple-to-process structures contribute a pragmatic/semantic interpretation to the dependency that is not present in gap dependencies. This pragmatic function, however, is not present in resumptive pronouns that occur in deeply embedded and island-violating dependencies. Hence, Erteschik-Shir (1992) distinguishes between two types of resumptive pronouns in Hebrew: grammatical resumptive pronouns, which occur in short-distance

dependencies to fulfil a pragmatic/semantic function, and ungrammatical resumptive pronouns, which occur in difficult-to-process constructions (embedding and islands) and serve a processing function.

Similarly, Sterian (2016) argued that RPs inside island structures are intrusive pronouns used to rescue island-violating dependencies in Iraqi Arabic. However, in non-island contexts, Sterian (2016) made a further distinction between two types of RPs in Iraqi Arabic, depending on whether they are obligatory or optional. In optional cases, such as with Discourse-linked (*D-linked*) questions, Sterian (2016) argued that RPs are grammatically licensed but are not required by syntax; instead, these optional RPs have a pragmatic function, as they trigger a different interpretation of the dependency than gap dependencies, as explained in (20).

20. A. Possible answers to questions with quantifiers - gap

ya: mraya kull riḏḏʒa:l ʕazam____
which woman every man invited.3MS

'Which woman did every man invite ____ ?'

- a. Natural function answer: his sister
- b. Pair-list answer: Samer, Suha; Ahmad, Najwa; etc

B. Possible answers to questions with quantifiers - resumption

ya: mraya kull riḏḏʒa:l ʕazam=ha
which woman every man invited.3MS=3FS

'Which woman did every man invite [her] ?'

- a. Natural function answer: his sister
- *b. Pair-list answer: Samer, Suha; Ahmad, Najwa; etc

(Iraqi Arabic; Sterian, 2016, pp.212–213)

On the other hand, when Iraqi Arabic RPs are obligatory, as in case of relative clauses (21), resumptive pronouns are treated as a purely syntactic phenomenon, the appearance of which is required for syntactic purposes but contributes nothing to the interpretation of the dependency.

21. a. With Gap:

*il-ka:teb illyi šeft ____ eb-be:t Suha tša:n

the-writer whom saw.1S___ at-house Suha was

Khalil Jubran

Khalil Jubran

'The writer whom I saw_____ in Suha's house was Khalil Jubran.'

b. With RP:

il-ka:teb illyi šeft=**ah** eb-be:t Suha tša:n

the-writer whom saw.1S=**him** at-house Suha was

Khalil Jubran

Khalil Jubran

'The writer whom I saw [**him**] in Suha's house was Khalil Jubran.'

(Iraqi Arabic; Sterian, 2016)

These accounts did not assume a distinction between the grammatical status of RPs inside island structures in true and intrusive resumption languages: both are ungrammatical but have a processing 'repair' function. True and intrusive resumption languages differ only in terms of whether or not resumptive pronouns are grammatically licensed in non-island contexts. Furthermore, these accounts appeal to the semantic and pragmatic features of pronouns to account for cases in which RPs are optional in true resumption languages.

This section has reviewed the theoretical and empirical aspects of the typology and variation of resumptive pronouns. In the theoretical literature, resumption has been treated as a non-uniform phenomenon both across languages or across constructions within the same language. Resumption has also been treated as a universal phenomenon, which performs a last resort function that is either syntactic or processing-related in nature.

The next section will focus on the experimental literature regarding resumption. This experimental literature is primarily concerned with investigating the extent to which the use of RPs in a certain language is grammatically licensed, as well as the extent to which resumptive pronouns in a certain language are sensitive to performance constraints. The first part reviews the experimental literature on the English language, while the second part reviews the experimental work on Hebrew and Arabic varieties.

2.2 Resumption in the experimental literature

2.2.1 Experimental literature on resumption in English

Studies in the theoretical literature contend that, in English, examples like (23) are more acceptable than (22):

22. *Who did Mary meet the people that will fire ___ ? (wh-movement from relative clause island)
23. Who₁ did Mary meet the people that will fire him₁? (resumptive pronoun inside relative clause)

(Phillips et al., 2019, p.6)

In the majority of cases, however, experimental acceptability judgement studies did not detect the improvement effect of resumption reported in the theoretical literature (Heestand et al., 2011; Keffala, 2011; Polinsky et al., 2013). In essence, intrusive RPs in English were consistently rated significantly lower than grammatical controls in acceptability judgement studies. According to these results, such intrusive RPs neither result in full acceptability nor improve the acceptability of complex-to-process dependencies in comparison to their gapped counterparts, disconfirming standard assumptions about intrusive pronouns in English. These findings have been taken as evidence to support the claim that intrusive RPs in English are ungrammatical.

However, Goodall (2017) provides a different interpretation for these findings. Observing that the acceptability of RPs in English remains low regardless of whether or not they appear in islands, while the acceptability of gaps is conditioned by the structure in which they occur, Goodall (2017) argues that filler-resumptive dependencies in English are grammatically licensed, but inherently ambiguous: at the point of parsing, they could be processed similarly to a gap, to resolve the dependency, or they could be perceived as a normal pronoun referring to some other discourse entity. Goodall (2017) argues that this ambiguity associated with interpreting RPs increases the processing load, resulting in poor acceptability. To account for the difference between true and intrusive RPs, Goodall (2017) claims that 'a whole range of properties might make resumptives more or less acceptable in any given language, even though the fact that they are allowed by the grammar remains constant across languages' (2017, p.74).

A different approach is advocated by Phillips et al. (2019), who attribute the mismatch between the claims of the theoretical literature and results of the experimental literature to methodological issues. These authors contend that the lack of any processing effect of RPs on complex-to-process dependencies might be an artifact of using a scalar acceptability judgement in experiments. In other words, Phillips et al. (2019) find it unlikely that linguists' introspective judgements of RP dependencies in English are mistaken. As noted by Ackerman et al. (2018), in acceptability judgement experiments, ungrammatical English island-violating dependencies with or without an RP may not be distinguishable by naïve participants presented with fully grammatical experimental sentences.

In fact, the mismatch between the claims of the theoretical and experimental research on the topic of resumption disappears when different methodological approaches are adopted. Ackerman et al. (2018) proposed that the facilitation effect of RPs is more likely to be evident when using methods that are both highly sensitive to acceptability and similar to methods used in formal syntax. Ackerman et al. (2018) investigated the amelioration effect of resumptive pronouns in island-violating sentences in a forced choice task, which asked participants to compare a minimal pair of sentences that differ only in the type of element at the dependency tail (gap vs RP). Ackerman et al. (2018) argued that this methodology is highly sensitive to acceptability, as it directs participants' attention to the exact part of the sentence that differs between the presented minimal pairs. Results revealed significant and robust preferences for RP dependencies over gap dependencies in island-violating dependency contexts.

The processing advantage of intrusive RPs over gaps in complex-to-process dependencies in English was further confirmed in psycholinguistic research using different methodological approaches. For instance, Beltrama and Xiang (2016) found that ratings of RP dependencies increase relative to those of gap dependencies when participants are asked to judge the comprehensibility of a sentence rather than its acceptability. Chacón (2019) further reports that resumption becomes more acceptable when participants are engaged in performing working memory (WM)-consuming tasks while judging the acceptability of experimental sentences. In online studies, resumption was found to speed up RT in dependencies involving multiple levels of embedding (Hofmeister and Norcliffe, 2013), as well as in dependencies that resolve inside island structures (Hammerly, 2019). The results of these studies challenge Goodall's (2017) view that the low acceptability of RPs is due to processing costs.

More detailed explanations of these studies will be presented in Chapter 3, which is dedicated to reviewing the factors that affect the distribution of RPs across languages. For now, it is sufficient for us to clarify the standard view concerning intrusive RPs in English: namely, that RPs are not grammatically licensed, but are utilised by language users when resolving FGDs that are too demanding (either due to an increase in processing demands or violation of a grammatical constraint).

2.2.2 Experimental literature on resumption in Irish, Hebrew and Arabic

Recent experimental investigations on resumption in true resumption languages are not fully consistent with the views of the theoretical literature. Corpus-based studies in true resumption languages have revealed that resumption is largely dispreferred in situations where processing demands are not high. McCloskey (2017, p.88) re-examined the distributional pattern of RPs in Irish corpus data and noticed that 'resumption is massively disfavored in usage'. At the beginning of the article, McCloskey said:

'I have claimed that the grammar of Irish makes available a free choice between resumption and filler-gap dependencies and that the choices actually made reflect performance factors. That is probably correct, but I have to confess that I have been shocked to discover, in looking more closely at the facts, how extreme the prejudice against resumption is' (2017, p.88).

Despite the fact that the theoretical literature on Irish resumption suggests that FGDs that terminate in non-island contexts allow for a free alternation between gaps and resumptive pronouns, McCloskey (2017) found that resumptive pronouns are strongly dispreferred compared with gaps in this context (8.8% percentage for RPs usage and 91.2% for gaps usage). The percentage of RP usage in comparison to gap usage improved 'at points of heightened processing pressure' (McCloskey, 2017, p.94). However, only resumptive pronouns rather than gaps appear in sentences that involve a violation of island constraints.

Similarly, Ariel's (1999) corpus-based study of resumption in Hebrew revealed that gaps are favoured over resumptive pronouns in non-embedded positions (10% resumption, 90% gaps). The proportion of RP usage in comparison to gap usage increased in more deeply embedded positions. More specifically, the usage of resumptive pronouns increased in comparison to gaps as the complexity of processing FGDs increased: RPs were used in 41.2% of cases in prepositional object

relative clauses compared to 8.7% of cases in object relative clauses. Notably, Ariel's corpus study does not include island-violating dependencies.

Experimental acceptability judgements were also not consistent with the claims in the theoretical literature regarding resumption in true resumption languages. In particular, experimental acceptability judgements fail to confirm the claim that true RPs are freely interchangeable with gaps in non-island contexts, or the claim that true RPs are fully grammatical inside islands.

Farby et al. (2010) investigated resumption behaviour in Hebrew in both island and non-island contexts in acceptability judgement experiments. A sample of their experimental sentences across conditions is presented in (24). Results revealed that resumptive pronouns significantly improve the acceptability of sentences containing island structures ($t(187)=3.41$, $p<0.01$), but that this improvement is negligible, since resumptive pronoun dependencies are still rated very low (2.31 on a five-point scale). In non-island contexts, gaps are found to be significantly preferred to resumptive pronouns ($t(187)=5.57$, $p<0.001$), although RPs are still rated as acceptable' (gaps: 4.25, RPs: 3.5 on a five-point scale).

24. a. Non-island, gap/resumption:

dina	maskima		lifgoš	et	ha-calemet
Dina	agrees		to	meet	the photographer
še	dan	pagaš	_/ota	be-xeyfa	be-mikre.
that	Dan	met	_/her in Haifa		by chance.

b. Island, gap/resumption:

dina	maskima	lifgoš	et	ha-calemet	
Dina	agrees	to	meet	the photographer	
še	ha-xaver	še	pagaš	_/ota	be-xeyfa nasa le-šam
that	the friend that met	_/her in Haifa	went	there	
					be-mikre.
					by chance.

(Hebrew; Farby et al., 2010)

Similarly, Meltzer-Asscher et al. (2015) investigated the acceptability of RPs in Hebrew non-island relative clauses. Experimental sentences were presented in both

written and auditory forms. Results revealed that direct object resumptive pronouns are consistently judged as less acceptable than gaps, suggesting that ‘the alternation between gaps and RPs is not completely free in Hebrew’ (Meltzer-Asscher et al., 2015, p.71).

Furthermore, the claim in the theoretical literature that true resumptive pronouns are perfectly acceptable was challenged by the results of an acceptability judgement study drawing on data from Modern Standard Arabic (Tucker et al., 2019). These findings indicate that the occurrence of resumptive pronouns inside islands does not make sentences grammatical; while their acceptability improved significantly, RPs still received very low ratings. As such, the ameliorative effect of resumptive pronouns in Modern Standard Arabic is described as ‘making the best of a bad situation’ (Tucker et al., 2019, p.37). These authors conclude that ‘in certain grammatical corners, a grammaticalised resumption language can behave like an intrusive resumption language in penalising the presence of a pronoun’ (2019, p.36).

To conclude, the experimental literature on true resumption languages revealed two results that are, surprisingly, not consistent with the claims of the theoretical literature. First of all, corpus-based studies suggest that the usage of RPs in true resumption languages is strongly disfavoured compared to gaps, as well as that the use of true RPs increases in line with processing demands. Second, acceptability judgement experiments suggest that true resumptive pronouns are not fully acceptable either inside or outside islands.

However, it is important to highlight that experimental investigation of true resumption languages tends to primarily investigate (i) the factors underlying the choice between gap and RP strategies, in cases where both are allowed, and (ii) the extent to which RPs repair islands. These studies tend not to test how resumption behaves in dependencies where (it is claimed that) RPs are obligatory, such as relative clauses in Arabic varieties (25) or Somali *wh*-questions (26).

25. a. With Gap:

*I-ktab illi štarayt___ mbariḥ ḍaḥa
the-book.M which bought.1S___ yesterday lost
‘The book that I bought___ yesterday is lost.’

b. With RP:

I-ktab illi štarayt-uh mbariḥ ḍaḥa

the=book.M which bought.1S=him yesterday lost
'The book that I bought [it] yesterday is lost.'

(Jordanian Arabic; Guillot and Malkawi, 2009)

26. a. **[Nin-kee]_F** b=aad sheegtay in=*(**uu**)
man-which FOC=2SG report.PST.2SG C=3SG.M
ku caayey?

2SG insult.PST.3SG

'Which man did you say insulted you?'

(Somali; Hedding, 2014)

- b. **[Kum]_F=ay** u maleynaysaa in=*(**uu**) Amina
who.FOC=she to think that=3SG.M Amina
arkay?

see.PST.3SG.M

'Who does she think saw Amina?'

(Somali; Saeed, 1984, p.144)

- c. **[Max]_F=aad** doonaysaa in=*(**ay**)
what.FOC=2SG want.PROG.2SG C=3PL
dhacaan?

happen.PRES.3PL

'What do you want to happen?'

(Somali; Saeed, 1984, p.144)

In fact, Friedmann and Costa's (2011) experimental study suggests that the grammatical status of obligatory RPs might differ from the grammatical status of optional RPs in true resumption languages. These authors investigated the comprehension of resumptive pronouns among two groups of hearing-impaired children and adolescents, who are known to have difficulties with processing long-distance dependencies. Participants were speakers of Hebrew and Palestinian Arabic; while both of these languages belong to the category of true resumption languages, the theoretical literature suggests that the use of RPs in Hebrew relative

clauses is optional, while it is obligatory in Palestinian Arabic relative clauses. A picture-matching task was used to measure performance on the comprehension of subject and object relative clauses against the performance of age-matched control groups.

The Hebrew-speaking, hearing-impaired participants were found to comprehend object relatives with RPs significantly better than their gapped counterparts. This finding is consistent with the hypothesis that resumption is preferred under performance pressures. On the other hand, the Palestinian Arabic-speaking hearing-impaired participants were shown to obtain no benefit from the presence of resumptive pronouns, as their comprehension of relative clauses in general was poor.

Friedmann and Costa (2011) accounted for the difference between the behaviour of impaired Hebrew and Palestinian Arabic speakers by arguing that resumptives differ between the two languages. Resumptive pronouns in Palestinian Arabic are clitics, which need to be licensed by a functional head; thus, they enter the derivation pre-syntactically as part of the numeration. Resumptive pronouns in Hebrew, on the other hand, are full pronouns that enter the derivation post-syntactically to facilitate the establishment of the dependency in relative clauses. While the findings of this study are consistent with theoretical accounts that distinguish between true RPs and intrusive RPs, they are inconsistent with the assumption that the Hebrew and Palestinian Arabic languages belong to the same class of true resumption languages.

In summary, the findings of experimental investigations into resumption in true resumption languages contradict the pervasive view articulated in the theoretical literature: namely, that true resumptive pronouns and gaps alternate freely in positions where only gaps are expected to occur in languages like English, and that true RPs are grammatically licensed in positions where gaps are not allowed inside island structures (Sells, 1984; Shlonsky, 1992; McCloskey, 2006). Although true resumptive pronouns in non-island contexts are acceptable, they are still not as acceptable as gaps in acceptability judgement experiments and are moreover highly dispreferred compared to gaps in corpus-based studies. The effect of resumption is still controversial inside islands; resumptive pronouns do not make island-violating dependencies grammatical, as suggested by the theoretical literature; island-violating dependencies with RPs are still perceived as unacceptable by participants, despite their marginal improvement effect.

The contrast in the processing of resumption in Palestinian Arabic and Hebrew by hearing-impaired children further complicates the picture. This observation, in

combination with the findings of corpus-based and acceptability judgement experiments in true resumption languages, leads us to concur with McCloskey (2017, p.95) that ‘the facts are richer and more subtle than a clear-cut distinction between “true” resumption and “intrusive” resumption’.

To conclude this section, we note that while the theoretical literature provides convincing proof that the two types of resumption exhibit distinct behaviour, the exact nature and extent of this difference remain unclear in the experimental literature; it is still not certain whether they represent the same phenomenon or two distinct phenomena. According to the former perspective, resumptive pronouns across languages are used either as last resort syntactic or last resort processing devices (Shlonsky, 1992; Ariel, 1999; Polinsky et al., 2013; Beltrama and Xiang, 2016). According to the latter view, true RPs represent a purely syntactic phenomenon, where resumption constitutes a grammatical mechanism for establishing A-bar dependencies, while intrusive resumption represents a processing (repair) phenomenon that is used as a rescuing device when it is difficult or impossible to establish FGDs (Sells, 1984; McCloskey, 2006; Sterian, 2016).

2.3 Language under investigation

The present study will focus on investigating the distributional properties of resumptive pronouns in different contexts in Baha Arabic, which I will argue belongs to the category of true resumption languages. I will thus adopt an experimental approach and attempt to add data to the literature on resumption; specifically, regarding the extent to which resumption in a true resumption language is grammatically licensed and the extent to which the distributional properties of RPs in a true resumption language differ from the distributional properties of RPs in intrusive resumption languages.

Varieties of Arabic, including Baha Arabic, make productive use of resumption across different types of dependency structures, making them an interesting environment for investigating the resumption phenomenon. It is therefore relevant to present a description of the different strategies used in Arabic varieties to form *wh*-questions and relative clauses. I will begin by describing some general properties of Baha Arabic to familiarise the reader with the language under investigation.

2.3.1 Basic description of Baha Arabic morphosyntax

Baha Arabic is the variety of Arabic spoken in the southern province of Saudi Arabia. As in any other part of the Arab world, Baha Arabic is a spoken variety of Arabic used in everyday communication, while Modern Standard Arabic (henceforth MSA) is the variety used in formal contexts (e.g. education, media, religious ceremonies, political speeches, etc.). The next section will describe the key properties that Baha Arabic has in common with MSA, as well as the ways in which they differ.

2.3.1.1 Word order

The word orders SVO and VSO are assumed to be unmarked word orders in MSA, while other word orders (OSV, OVS, SOV, VOS) are often used when the object is contrastively focused (Fassi-Fehri, 1993; Mohammad, 2000; Aoun et al., 2009). Similarly, all these different word orders are common in Baha Arabic (27); however, like other spoken varieties of Arabic, including Palestinian (Mohammad, 2000) and Moroccan (Fassi-Fehri, 1993; Aoun et al., 2009), the unmarked word order is SVO.

27. a. Mhammad ?kal at-tuffaḥah (Unmarked) SVO
'Mhammad ate the-apple.
- b. ?kal Mhammad at-tuffaḥah VSO
Ate Mhammad the-apple
- c. at-tuffaḥah Mhammad ?kal OSV
The-apple Mhammad ate
- d. at-tuffaḥah ?kal Mhammad OVS
The-apple ate Muhammad

2.3.1.2 Agreement

MSA exhibits an agreement asymmetry between the subject and the verb, depending on the word order used. While the verb fully agrees in terms of gender, person and number with the subject in SVO order (see (28)), it only partially agrees with it (in relation to gender) in VSO order (see (29)).

28. a. al-?wlad-u šahad-uu al-mudarris-a (SVO)
The-boys.NOM saw.3MP the-teacher-MS.ACC

wrote-3MS Saeed-NOM the-lesson-ACC

'Saeed wrote the lesson.'

32. a. Al-banat-u al-jameelat-u
 The-girls-NOM the-beautiful-NOM
 The beautiful girls.'
- b. *Al-banat-u al-jameelat-i
 The-girls-NOM the-beautiful-GEN
 'The beautiful girls.'

Moving to Baha Arabic, nouns, verbs and adjectives are not case-marked, as illustrated in (33).

33. a. Al-bent katabat ad-dars
 The-girl wrote the-lesson
- b. Al-banat al-ḥelwat
 The-girls the-beautiful
 'The beautiful girls.'

2.3.2 Wh-formation and relativisation strategies in Arabic varieties

In terms of *wh*-questions, Arabic varieties, including Baha Arabic, have three distinct strategies for forming *wh*-questions: *wh*-gap, *wh*-resumption, and *wh*-cleft.

The gap strategy is the default strategy for forming *wh*-questions in the majority of Arabic varieties. The hallmarks of this strategy are the fronting of a *wh*-phrase, the absence of the relative complementiser *illi*, and the appearance of a gap at the extraction site. The following examples from different varieties of Arabic illustrate this strategy.

34. ʔay bant šaft-ii____?
 which girl see-2FSG
 'Which girl did you see?'

(Saudi Arabic; Alshaalan and Abels, 2020)

35. [šuj] ?al-l-ek iyad tʕml-i-l-u t;
what said-to-you lyad make-2SG.F.SU-for-3SGM.OBJ
ʕa-l-ʕasha?
on-the-dinner
'What did lyad ask you to make for him for dinner?'
(Syrian Arabic; Sulaiman, 2016)

The resumptive strategy is similar to the gap strategy, except that a resumptive pronoun rather than a gap appears at the dependency tail.

36. ?ay bant šaft-ii-hā?
which girl see-2FSG=her
'Which girl did you see her?'
(Saudi Arabic; Alshaalan and Abels, 2020)
37. miin/?ayya maariḍ zarit-u naadia?
who/which patient visited-3SG.F=her Nadia
'Who/which patient did Nadia visit?'
(Syrian Arabic; Sulaiman, 2016)

The literature on *wh*-questions in Arabic varieties (Aoun et al. 2010) suggests that inanimate bare *wh*-phrases (e.g. 'what') are excluded from the resumptive strategy (38). However, this conclusion seems not to apply to varieties such as Saudi Arabic and Jordanian Arabic. The examples below show that resumption is possible with inanimate bare *wh*-phrases in Jordanian Arabic (39) and Saudi Arabic (40).

38. *šū štarayt-i
what bought.2SF=it
'What did you buy?'
(Lebanese Arabic; Abdel Razaq, 2011)
39. eyš tjarribt-i b-l-?awwal?
what try.2FS-it in-the-first
'What did you try first?'

(Jordanian Arabic; Abdel Razaq, 2011)

40. ʔayš Nawf akala-t men-uh?

what Noaf eat-3FSG from-it?

‘What did Noaf eat from?’

(Saudi Arabic; Alshaalan and Abels, 2020)

Moving on to the *wh*-clefting strategy, the hallmarks of this strategy are the fronting of a *wh*-phrase, the obligatory presence of the relative complementiser *illi*, and the presence of an obligatory overt resumptive pronoun. The following examples illustrate this strategy.

41. ʔay bant (hay) aly šaft-ii-hā?

which girl she that see-2FSG=her

‘Which girl is it that you saw?’

(Saudi Arabic; Alshaalan and Abels, 2020)

42. miin (hi) ʔilli l-ʔasad ʔakal-ha mbaarih?

who she that the-lion ate-her yesterday

‘Who did the lion eat yesterday?’

(Syrian Arabic; Sulaiman, 2016)

Similarly, definite relative clauses are characterised by the presence of a relative head noun phrase, an obligatory relative complementiser (*illi*), and an obligatory overt resumptive pronoun.

43. il-ka:teb illyi šeft=**ah** eb-be:t Suha tša:n Khalil

The-writer whom saw.1S=**him** at-house Suha was Khalil

Jubran

Jubran

‘The writer whom I saw [**him**] in Suha’s house was Khalil Jubran.’

(Iraqi Arabic; Sterian, 2016)

Indefinite relative clauses, on the other hand, lack the relative complementiser ‘*illi*’, but obligate the presence of an overt resumptive pronoun, as in (44).

44. ʕam fattiš ʕa kteeb (*yalli) Dayyaʕt-o l-yom

Asp look.1S for book that lost.1s-it today

'I am looking for a book that I lost today.'

(Aoun et al., 2009)

In this thesis I will deal only with definite relative clauses.

2.3.3 Resumption in Highest Subject position

An important note is in order regarding cases in which overt resumptive pronouns are optional or obligatory. A widely accepted generalisation in the literature on resumption in true resumption languages is that an overt resumptive is impossible in the highest subject position (45). This can be due either to the subject *pro*-drop status of a language (Sultan, 2007), or the Highest Subject Restriction (Shlonsky, 1992; McCloskey 2006). Hence, the present study will focus solely on investigating the behaviour of resumption in the object position.

45. *an fear a raibh sé breoite

the man COMP was [he] ill

'the man that was ill'

(McCloskey, 1990)

2.3.4 Pronominal status of resumptive pronouns in Arabic

According to Aoun et al. (2009), weak pronouns (i.e. clitics) are used as resumptive pronouns in all varieties of Arabic; resumptive clitics appear attached to a verb, a noun or a preposition in non-subject positions. See (46).

46. ?ayya mmasil šəft-o b-l-maTʕam?

which actor saw.2ms-him in-the-restaurant

'Which actor did you see in the restaurant?'

(Aoun et al., 2009, p.8)

Aoun et al. (2009), furthermore, noticed that there is variability in whether strong pronouns can be used as resumptives across varieties of Arabic; for instance, while strong pronouns cannot be used as resumptives in Moroccan Arabic (47), they can be used as resumptives in Lebanese Arabic (48).

47. a. šmen Talib nsiti fin tlaqiti-h (*huwwa)

which student forgot-2s where met-2s-him (HIM)

'Which student have you forgotten where you met?'

b. šmen Talib saferti qblma yTerdu-h (*huwwa)

which student traveled-2s before expelled-3p-him (HIM)

'Which student did you travel before they expelled?'

(Aoun et al., 2009, p.7)

48. a. ʔayya tilmiiz nsiit-e ween ltaʔay-te fi-i huwwa

which student forgot-2fs where met-2fs in-him HIM

'Which student did you forget where you met?'

b. ʔayya tilmiiz seefar-te ʔablma yišhaT-u -u huwwa

which student traveled-2fs before expelled-3p him (HIM)

'Which student did you travel before they expelled?'

(Aoun et al., 2009, p.7)

In this thesis I will deal only with wh-dependencies involving resumptive clitics.

2.4 Conclusion

This chapter has presented an overview of resumption typology, as described in the theoretical and experimental literature. A dominant view in the theoretical literature is that intrusive RPs and true RPs represent two distinct phenomena: specifically, true resumption is viewed as a grammatically licensed mechanism for creating long-distance dependencies in certain languages, while intrusive resumption is viewed as a rescuing device utilised when the processing demands of long-distance dependencies are very high. Some theories, however, do not assume a distinction between these two types of RPs in their representational properties. Instead, they consider resumptive pronouns in both true and intrusive resumption languages to have the same last resort processing/syntactic function.

Baha Arabic was further introduced as the language under investigation. We believe that this language, as an under-studied variety of Arabic, is an interesting choice for investigations of resumption, since it features a productive use of resumption across different types of dependency structures. The literature on resumption in Arabic varieties has mainly focused on formal syntax, except for Tucker et al.'s (2019)

experimental study. As illustrated above, the results of this study were not consistent with the predictions of the theoretical literature, as the acceptability of resumptive pronouns was found to be highly degraded. However, a notable shortcoming of this study is that participants' proficiency in MSA could potentially have been insufficient to allow them to take part in acceptability judgement experiments of this kind. In other words, it is frequently claimed that there are no 'native speakers' of MSA (Kaye, 1970; Maamouri, 1998), since the use of MSA is restricted to formal occasions.

In the present study, I aim to contribute to the current literature on resumption by investigating resumption in a spoken variety of Arabic. The fact that resumption use is productive in this language allows us to investigate resumption behaviour in a wide range of types of dependency structures; as illustrated in the theoretical literature, some of these obligatorily require RPs, others optionally allow RPs, and still others do not allow RPs at all. The next chapter is dedicated to reviewing the syntactic and non-syntactic factors that contribute to the distributional properties of RPs across languages, as proposed in the syntactic and psycholinguistic literature.

Chapter 3 Resumptive Pronouns in Syntax and Psycholinguistics

3.1 Introduction

In this chapter, we will review the syntactic and psycholinguistic accounts of the role of RPs in syntax and the processing of FGDs. Generally speaking, it is widely accepted that the use of resumption is limited to cases where locality constraints on FGDs are applied; in other words, that usage of RPs is restricted to cases where gaps are either dispreferred or not allowed to occur.

Hence, to understand the syntactic and psycholinguistic accounts of resumption, this chapter will begin by explaining the locality constraints on the formation of FGDs. We will then review the syntactic theories of the interaction between locality constraints and resumption. The assumptions of two such theories will be reviewed: the Last Resort account and the C(omplementiser)-type account. The former attempts to explain the distinction between gap dependencies and RP dependencies in syntactic terms, but assumes no distinction between true and intrusive RPs; the latter claims that a distinction should be made between true and intrusive RPs and consequently attempts to identify the factors that cause the grammar of true resumption languages to be different from the grammar of intrusive resumption languages. In essence, the two accounts differ in that the former assumes true and intrusive RPs to have the same underlying syntactic structure (which differs from that of gapped dependencies), while the latter assumes a distinct underlying syntactic structure for true RP dependencies on one hand and a distinct underlying syntactic structure for gap and intrusive RP dependencies on the other hand.

Next, we will review a processing theory that attempts to explain locality constraints in terms of constraints on working memory capacity: namely, the Resource Limitation Theory, or RLT (Kluender and Kutas, 1993a; Kluender, 1998). We will further introduce a WM-based account of the interaction between locality constraints (as described in RLT) and resumption. Importantly, these theories assume that there is no difference in the underlying syntactic structure between gap dependencies and (true/intrusive) RP dependencies.

Consequently, we will argue that the empirical facts and experimental findings pertaining to the variation in resumption across languages and structures are best explained by a theory that assumes a distinct underlying syntactic structure for dependencies created with true RPs and those created with gaps and intrusive RPs.

At the same time, WM-based theories will be shown to be essential to explaining the distributional properties of RPs within languages and structures.

After reviewing the relevant literature, we will introduce the diagnostics that we will initially adopt to distinguish between true and intrusive RPs in Baha Arabic, the language under investigation. Finally, we will address methodological concerns regarding the use of acceptability judgement methods in the study of linguistic phenomena, particularly those located at the interface between syntax and processing.

3.2 Background on locality constraints

Despite the fact that FGDs are able to cross an indefinite number of clauses, the linguistic and psycholinguistic literature identifies positions where gaps are possible but dispreferred (1), as well as positions where gaps are not allowed, as in (2) and (3):

1. a. This is the girl_i that John likes ___i
- b. This is the girl that Peter said that John likes (gap).
- c. This is the girl that Peter said that John thinks that Bob likes (gap).
- d. ?This is the girl_i that Peter said that John thinks that yesterday his mother had given some cakes to ___i.

(modified from Erteschik-Shir, 1992)

2. *I just saw a girl who Long John's claim that ____ was a Venusian made all the headlines.

(Ross, 1986, p.260)

3. *This is the man whom_i Emsworth told me [when he will invite ___i].

(Haegeman, 1994)

The effect illustrated in (1) is referred to as an 'embedding effect': i.e. FGDs are worse when several clause boundaries intervene between the filler and the gap. The effect illustrated in (2) and (3) is known as an 'island effect': gaps are not allowed to occur in certain syntactic configurations, metaphorically known as 'islands'.

A dominant hypothesis in the linguistic literature assumes that the distributional pattern of RPs across languages is sensitive to these locality constraints. In other

words, RP dependencies are preferred to gap dependencies when locality constraints are violated (Ross, 1967; Shlonsky, 1992; Erteschik-Shir, 1992; Ariel, 1999).

For instance, Erteschik-Shir (1992) argued that the complexity associated with processing embedded FGDs is ameliorated when the embedded gap is replaced with a resumptive pronoun. The deeply embedded gap dependencies in (1) are repeated in (4) below with resumptive pronouns. Erteschik-Shir (1992) argued that the acceptability of RPs increases as the level of embedding also increases.

4. a. This is the girl that John likes t/*her
- b. This is the girl that Peter said that John likes t/??her
- c. This is the girl that Peter said that John thinks that Bob likes t/?her
- d. ?This is the girl that Peter said that John thinks that yesterday his mother had given some cakes to ?t/her

(Erteschik-Shir, 1992)

Furthermore, resumptive pronouns have been argued to have an amelioration effect on islands across languages. For instance, the island-violating dependency in (2) is repeated below in (5), first with a gap and then with an RP. We can observe that the acceptability of the island-violating dependency increases when an RP ('she') occurs at the dependency tail rather than a gap.

5. a. *I just saw a girl who Long John's claim that ____ was a Venusian made all the headlines.
- b. I just saw a girl who Long John's claim that she was a Venusian made all the headlines.

(Ross, 1986, p.260)

The pervasive view in the linguistic literature is that locality constraints resulting in embedding effects represent constraints on real-time sentence processing, as they identify positions where gaps are possible (i.e. allowed by the grammar) but dis-preferred; in other words, dependencies with deeply embedded integration sites are grammatical, but consume large amounts of WM resources (Kluender and Kutas, 1993a; Kluender, 1998). On the other hand, the dominant view of the underlying conditions of locality constraints that result in island effects assumes that these constraints represent grammatical constraints that identify structures where gaps are prohibited; i.e. island effects arise due to the violation of grammatical constraints that

prohibit gaps from occurring inside island structures (Ross, 1967; Chomsky, 1977; Huang, 1982; Chomsky, 1986; Rizzi, 1990).

However, recent linguistic and psycholinguistic literature on FGDs has started debates as to whether all these locality constraints – those resulting in island effects and those resulting in embedding effects – are reducible to independently motivated constraints on linguistic processing, or whether a distinction between them must be maintained (Deane, 1991; Kluender and Kutas, 1993b; Kluender, 1998; Hofmeister and Sag, 2010; Goodall, 2014). As a result, the source of island effects remains a controversial topic, which complicates accounts of their interaction with resumption.

In the next section, we will review the syntactic accounts of the interaction between locality constraints resulting from the island effect and resumption; in essence, these accounts view islands as representing grammatical constraints on the formation of FGDs, and further argue that the distinction between RP dependencies and gap dependencies is fundamentally grammatical in nature.

3.3 Syntactic accounts

3.3.1 FGD formation in syntax

To derive a syntactic dependency as in (6) below, the syntactic literature assumes that the filler phrase ‘what’ originates in its canonical position after the verb ‘wrote’ and moves to the initial position in the sentence (i.e. the specifier of the complementiser phrase (CP)) to satisfy features of C, leaving a trace behind at the gap site.

6. [_{CP} What_i does John think [_{CP} that Jane wrote ____i]]?

However, this movement must operate in a successive-cyclic mode. For instance, in (6), the filler phrase ‘what’ can only move to its matrix clause position if it stops at the specifier position of the embedded clause (spec-CP). That is, movement does not take place in one single step, but rather must proceed in a successive-cyclic fashion. In their original formulations, such constraints include the Subjacency Condition (Chomsky, 1973), which contends that movement may cross at most one bounding node at a time, where bounding nodes are IP and DP (as in (7) and (8)), and condition on extraction domains (Huang, 1982), which states that *wh*-movement is forbidden from non-complements (as in (9)).

7. [_{CP} What_i did [_{IP} Bill claim [_{CP} that [_{IP} he read ____i]]]] ?

8. *_{[CP} What_i did _{[IP} Bill make _{[DP} the claim _{[CP} that _{[IP} he read ___ _{i]]]]} ?
9. *_{[CP} Who_i did Mary cry _{[PP} after John hit ___ _{i]]}?

(Huang, 1982)

Chomsky (2001) proposed that syntactic operations occur in small chunks of structures, referred to as 'phases'. In its original form, the syntactic objects that form phases are defined as vP and CP. According to this theory, syntactic successive-cyclic movement is attributed to the so-called Phase Impenetrability Condition, which ensures that only the edge of the phase and its head are accessible to syntactic operations associated with higher phases; that is, the phase impenetrability condition forces movement to operate successive cyclically via phase edges. Movement is only permitted if a phrase has first moved to the left edge of its respective phase.

3.3.2 Islands in syntax

Ross (1967) observed that FGDs cannot be established across some structures, and gave these the metaphorical term 'islands'. In particular, a filler phrase cannot establish a dependency with another element located within these islands; otherwise, an 'island effect' will emerge. The different types of island identified in literature are illustrated in (10) below.

10. a. Relative Clause Island (Complex NP):
* **Who** did Dale comfort [NP the woman that [S saw ___ ?]]
- b. Whether Island:
* **Who** did Dale wonder [whether Bob frightened ___ ?]
- c. *Wh*-Island:
* **Who** did Dale say [who saw ___ behind Laura's bed?]
- d. Subject Island:
* **Who** did [the fact that Sarah saw ___] surprise Dale?
- e. Adjunct Island:
* **Who** did Dale ruminate [while Harry interrogated ___ ?]

f. Coordinate Structure Constraint:

* **Who** did [Dale suspect ____ and Harry interrogate Leland?]

g. Factive Island:

* **Why** did Dale remember [that Ben was suspicious ____?]

(Chacón et al., 2016, pp.3-4)

The dominant view of island effects is that the syntactic successive-cyclic movement is blocked when gaps occur inside island structures (Ross, 1967; Chomsky, 1973; Chomsky, 1986; Rizzi, 1990). In other words, syntactic accounts claim that island effects emerge when a well-formed successive-cyclic movement is not available, whether due to a violation of the subjacency condition (Chomsky, 1973; Chomsky, 1977), the condition on extraction domains (Huang, 1982), or the Phase Impenetrability Condition (Chomsky, 2001).

Experimental investigations of FGDs that cross island boundaries have revealed that extraction from these structures usually results in an extreme degradation in acceptability; this degradation is described as robust and consistent (Alexopoulou and Keller, 2007; Heestand et al., 2011; Goodall, 2014; Michel, 2014; Kush et al., 2015; Sprouse et al., 2016; Tucker et al., 2019).

3.3.3 Resumption in syntax

Syntacticians claim that replacing a gap inside islands with an RP can ameliorate the acceptability of island-violated dependencies (Ross, 1967; Chomsky, 1977). In the following examples, we can observe that the presence of the pronominal element at the dependency tail seems to ameliorate violation of the island constraint in English.

11. a. *I just saw a girl who Long John's claim that ____ was a Venusian made all the headlines.

b. I just saw a girl who Long John's claim that she was a Venusian made all the headlines.

(Ross, 1986, p.260)

12. a. This is the man whom_i Emsworth told me when he will invite him_i.

b. This is the man whom_i Emsworth made the claim that he will invite him_i.

(Haegeman, 1994)

Merchant (2004) offers the following example from Greek, which further illustrates the rescuing effect of RPs on island-violating sentences:

13. Janis ine o adras pu i
 Janis.NOM is the.NOM man.NOM that the
 Maria efiWe apo o parti otan
 Maria left.3SG from the party when
 ton iLe
 hi saw.3SG

'Janis is the man Maria left from the party when she saw him.'

(Greek; Merchant, 2004)

The 'saving' effect of resumptive pronouns on island-violating sentences can be observed across the different varieties of Arabic. The examples (14), (15) and (16) below present island-violating examples from different Arabic varieties. We can observe that the island-violating dependencies with the resumptive pronoun are deemed acceptable; however, lack of a pronoun renders these sentences unacceptable. In other words, the resumptive pronoun seems to rescue these island-violating dependencies.

14. a. Sməʕt ʔənnu Naadya, byaʕrfo ʔayya walad
 heard.1S that Nadia know.3P which boy
 jeef=a
 saw.3S=her

'I heard that Nadia, they know which boy saw her.'

- b. * Sməʕt ʔənnu Naadya, byaʕrfo ʔayya
 heard.1S that Nadia know.3P which
 walad jeef=__
 boy saw.3S=__

'I heard that Nadia, they know which boy saw __.'

(Lebanese Arabic; Aoun and Benmamoun, 1998)

15. a. ʕa:lib-ha l-kasul ma bidku txabbro

student.M=hers the-bad Neg want.2PI tell.2PI
wala mʕallmih ʕan I-bent illi saʕadat=uh
no teacher.F about the-girl which helped=him
b-l-faḥṣ
in-the-exam

'Her bad student, you don't want to tell any teacher about the girl who helped him'

b. *ʕa:lib-ha I-kasul ma bidku txabbro
student.M=hers the-bad Neg want.2PI tell.2PI
wala mʕallmih ʕan I-bent illi saʕadat=__
no teacher.F about the-girl which helped=__
b-l-faḥiṣ
in-the-exam

'Her bad student, you don't want to tell any teacher about the girl who helped __'

(Jordanian Arabic; Guilliot and Malkawi, 2009)

16. a. ya: ʕu:ra.F li-bni=ha Samer ysʔal iða
which picture.F of-son=her Samer ask.3MS if
kull mraya šagagat=ha
every woman tore.3FS=3F

'Which photo of her son did Samer wonder if every woman tore [it]?'

b. * ya: ʕu:ra.F li-bni=ha Samer ysʔal iða
which picture.F of-son=her Samer ask.3MS if
kull mraya šagagat__
every woman tore__

'Which photo of her son did Samer wonder if every woman tore __?'

(Iraqi Arabic; Sterian, 2016)

Under the assumption that island effects emerge from the violation of constraints on syntactic movement, the alleviation of that effect is interpreted as an indication that no syntactic movement took place. This has led researchers to propose that there is no movement involved in the derivation of dependencies featuring resumptive pronouns, as they appear insensitive to island effects (Shlonsky, 1992; McCloskey, 2002; McCloskey, 2006). In this case, both the RP and the filler are assumed to be base-generated (i.e. externally merged from the lexicon), with a *binding* relation established between them. Generally speaking, syntactic accounts assume that gap dependencies are derived via successive-cyclic movement, while RP dependencies do not involve movement (Chomsky, 1977; Aoun et al., 2009).

Below, we will present the two dominant theories of the mechanisms involved in the choice between the gap strategy and the RP strategy.

3.3.3.1 Last resort or not last resort

As outlined in section 2.1.2, Shlonsky (1992) argued that the use of RPs across languages is conditioned by a Last Resort principle to rescue successive-cyclic movement operations. This view correctly accounts for the observation that RPs save or ameliorate islands cross-linguistically. However, cases in which the resumptive strategy and the gap strategy alternate freely in positions where movement is not blocked are problematic for the last resort theory. The following example from Irish illustrates that gap and resumptive pronouns are equally acceptable in the embedded object position.

17. a. an fear aL bhuaill tú ____
 the man COMP struck you ____
- b. an fear aN bhuaill tú é
 the man COMP struck you him
- ‘the man that you struck’

(Irish; McCloskey, 1990, p.18)

To account for such cases, Shlonsky (1992) argues that the choice between gaps and resumption in non-island contexts is conditioned by the choice of complementiser type. In essence, he assumed that some types of complementisers block successive-cyclic movement; when this occurs, resumptive pronouns are obligatorily used to

establish the syntactic dependency by means of binding relations. For instance, Shlonsky (1992) argued that use of the relative complementiser *illi* in Palestinian Arabic always makes movement impossible, which further illustrates why RPs are obligatory in Palestinian Arabic relative clauses, regardless of the presence of a crossed-island clause boundary (18). Shlonsky (1992) further extends this account to Hebrew, arguing that Hebrew features two, morphologically identical, complementisers *še*: one of them blocks movement, similar to the relative complementiser *illi* used in Palestinian Arabic, while the other one does not. (See (19)).

18. a. *l-bint ijilli šufti-____
the-girl that (you-f) saw-gap
- b. l-bint ijilli šufti-ha
the-girl that (you-f) saw-her
'the girl that you saw'
(Hebrew; Shlonsky, 1992, p.445)

19. a. ha-'iš še ra'iti ('oto)
the-man COMP saw-I (him).
- b. ha-'iš še ra'iti (____)
the-man COMP saw-I (____).
'The man that I saw'
(Hebrew; Shlonsky, 1992, p.444)

Shlonsky (1992) goes on to argue that the difference between languages like Hebrew and Palestinian Arabic and languages like English is that the former set of languages have more contexts in which movement is blocked than the latter. Specifically, in addition to island structures, there are other types of complementiser that block movement in Hebrew and Palestinian Arabic; English, for its part, lacks this type of complementiser.

Importantly, the last resort account draws a distinction between the derivation of gap dependencies and RP dependencies: specifically, unlike gap dependencies, movement is cancelled in the derivation of RP dependencies cross-linguistically. On

the other hand, no distinction is made between the derivational properties of RP dependencies in true and intrusive resumption languages. Hence, this theory attempts to treat resumption across languages as a universal phenomenon.

Rouveret (2011) criticised Shlonsky's (1992) last resort account, arguing that drawing a connection between the choice between gap and RP dependencies and the type of complementiser is not compatible with the syntactic last resort view of RPs. In other words, Rouveret (2011) argued that if a relative clause is introduced along with the complementiser in a way that blocks movement, resumptive pronouns are obligatorily selected; that is, they should not be thought of as last resort expressions. Hence, Rouveret (2011), claims that, 'If the gist of the analysis is simply that some complementisers are incompatible with movement, there is no need to resort to last resort' (p.12). This is the view adopted by the syntactic theories described below.

3.3.3.2 Complementiser type account

For the most part, the distributional pattern of RPs in syntactic literature has been investigated in connection with complementiser type (McCloskey, 2002; Alexopoulou, 2010). In essence, the authors argue that complementisers play a crucial role in the choice between the gap and the resumptive strategy. Cross-linguistically speaking, it can be observed that true RPs occur along with a special type of complementiser, distinct from the complementiser in gap dependencies. For instance, McCloskey (1990) noted that, in Irish, a correlation is observed between the presence and absence of resumptive pronouns and the form of the complementiser. While the resumptive strategy is used with the complementiser aN, the gap strategy is used with the complementiser aL. See (20).

20. a. [NP NP [CP aN [IP...pronoun...]]]
b. [NP NP [CP aL [IP...t...]]]
c. [NP NP [CP aN [IP...[CP go [IP...pronoun...]]]]
d. [NP NP [CP aL [IP ...[CP aL [IP...t...]]]]

(McCloskey, 1990)

Similarly, in Slavic languages, the choice between a gap and an RP in relative clauses appears to be sensitive to the structure of the CP layer. Lavine (2003) observed that

a fully inflected relative pronoun occurs with a gap, while a morphologically invariant complementiser occurs with an RP. See (21) and (22).

21. Ukrainian

a. Ce toj dim, jakyj
this the house:NOM.MASC.SG which:ACC.MASC.SG
ja bac=yv t vc-ora.
I:NOM saw yesterday

b. Ce toj dim, **éc=о** ja
this the house:NOM.MASC.SG COMP I:NOM
joho bac=yv vc-ora.
RP:ACC.MASC.SG saw yesterday

‘This is the house that I saw yesterday.’

(Lavine, 2003)

22. Polish

a. On spotkał tego chłopca,
he met the boy:ACC.MASC.SG
którego widziało t wczoraj.
whom:ACC.MASC.SG you-saw yesterday

b. On spotkał tego chłopca, **co**
he met the boy:ACC.MASC.SG COMP
go widziało wczoraj.
RP:ACC.MASC.SG you-saw yesterday

‘He met the boy that you saw yesterday.’

(Lavine, 2003)

In varieties of Arabic, resumptive pronouns are obligatory in both relative clauses (23) and clefted *wh*-questions (24); both structures are obligatorily introduced with the complementiser *illi* (Aoun et al., 2009; Abdel Razaq, 2011; Sulaiman, 2016; Alshaalan and Abels, 2020). Unlike the Irish and Slavic languages, which feature both a gap-

complementiser and RP-complementiser in relative clauses, relative clauses and clefted *wh*-questions in Arabic varieties lack a gap complementiser. This further supports the claim that RPs in Arabic relative clauses are obligatory in both island and non-island contexts.

23. Relative clauses

l-bint ?illi šufti-*(ha)
the-girl COMP saw-2SGF=her

'the girl that you saw.'

(Palestinian Arabic; Shlonsky, 1992)

24. Clefted *wh*-questions

a. miin ?illi l-?asad ?akal=ha mbaarih?
 who that the-lion ate=her yesterday

Who did the lion eat yesterday?

b. *miin ?illi l-?asad ?akal mbaarih?
 who that the-lion ate yesterday

Intended: 'Who did the lion eat yesterday?'

(Jordanian Arabic; Sulaiman, 2016)

It is thus unsurprising that syntactic accounts link the difference between the derivation of RP and gap dependencies to the different featural specifications of elements in the CP layer. Since RP dependencies are insensitive to islands, while gap dependencies are sensitive to islands, it is assumed that the alternation between types of complementiser manifests an alternation between movement and binding dependencies: specifically, RP complementisers initiate binding dependencies, while gap complementisers always trigger movement operations (Alexopoulou, 2010).

McCloskey (2002) provides an analysis of Irish data that relates the distributional properties of RPs with a featural specification of the elements in the CP layer. At the edge of each phase, C carries an EPP feature and (optionally) an Operator (*Op*) feature. An 'Agree relation' between C and an element in its c-command domain is triggered if C is associated with an *Op* feature. The Agree relation is followed by Move to satisfy the EPP feature. When C does not carry an *Op* feature, no Agree relation is triggered. In this case, the EPP feature of C is satisfied by merging the null operator

at spec-CP in its position, as a Merge is preferred over Move for reasons of economy. Then, a binding relation between the null merged operator at spec-CP and a pronominal is established at LF.

According to Alexopoulou (2010), the *Op* feature in McCloskey's (2002) proposal is a purely syntactic feature, the presence of which only affects the derivation of a dependency but has nothing to do with its interpretation. Hence, Alexopoulou (2010) argued that it is 'conceivable that languages vary with respect to the availability of such a feature on their relative C' (p.7). In some languages, C always lacks the *Op* feature, resulting in obligatory resumptive pronouns across all the different positions. This is the case in Lebanese Arabic and Palestinian Arabic relative clauses (see (25)).

25. a. l-bint ʔilli šufti-*(ha)
 the-girl COMP saw-2sg-fem-her
 'the girl that you saw'
- b. l-bint ʔilli fakkarti ʔinno
 the-girl COMP thought-2SG-F COMP
 *(hiy) raayh.a l-el-bait
 (she) going to-the-house
 'the girl that you thought that (she) is going home'
 (Shlonsky, 1992)

Languages such as Hebrew, Irish and Slavic optionally allow the presence of this feature on C in non-island contexts; when a dependency crosses an island boundary, C always lacks the *Op* feature. This explains the optionality of RPs in non-island contexts, as well as the obligatoriness of RPs in islands in these languages. See (26) and (27).

26. a. an ghirseach aL ghoid na sílogaí
 the girl COMP stole the fairies
 'the girl that the fairies stole away'
- b. an ghirseach aN-r ghoid na sílogaí í
 the girl COMP-[PAST] stole the fairies her
 'the girl that the fairies stole away'

(Irish; McCloskey, 2002)

27. a. ha-ʔiš še- raʔiti (ʔoto)
the-man that-(I) saw (him)
'the man that I saw'
- b. ha-ʔiš še- xa šavt še-(hu) melemed
the-man that-(you.F) thought that-(he) teaches
ʔanglit
English
'the man that you thought teaches English'

(Hebrew; Shlonsky, 1992)

Languages like English, on the other hand, only feature the type of C that bears the *Op* feature. Hence, in English, the *Op* feature on C always triggers movement. This explains the ungrammaticality of resumptive pronouns in English. For this reason, this syntactic theory assumes that the amelioration effect of resumptive pronouns in English cannot be described in terms of formal syntax.

Consequently, an important assumption of syntactic accounts bearing features of C, according to Alexopoulou (2010), is that RPs do not repair islands by themselves; rather, island-violating dependencies are sensitive to the type of the dependency (i.e. movement vs binding). The type of complementiser and its morpho-syntactic features determine the choice between the two types of dependencies. Hence, true resumptive pronouns do not save islands, but rather appear in binding dependencies as a consequence of the absence of movement. Intrusive RPs, on the other hand, do appear in movement dependencies. Alexopoulou (2010) consequently argued that, while true RPs in binding dependencies are expected to restore island-violating dependencies into full acceptability, intrusive RPs can ameliorate island effects but not fully repair them.

Importantly, these syntactic theories assume a distinction between the derivational properties of RPs in true resumption languages on one side and gap dependencies and RPs in intrusive resumption languages on the other side. While movement is involved in the derivation of intrusive RP dependencies and gap dependencies, it is not involved in the derivation of true RP dependencies.

3.4 WM-based accounts

The basic concept of processing-based accounts of locality constraints is that the degraded acceptability of deeply embedded or island-violating dependencies occurs due to the difficulty associated with the online processing of these structures. In other words, these accounts assume that the effect of locality constraints is motivated by restrictions on WM capacity (Deane, 1991; Kluender and Kutas, 1993b; Kluender, 1998; Hofmeister and Sag, 2010). Several accounts have been proposed in the literature to account for the mechanisms by which these constraints affect the processing of FGDs. Below, we will present the resource-limitation theory (Kluender and Kutas, 1993a; Kluender and Kutas, 1993b; Kluender, 1998), as it represents the most detailed account of the complexity associated with processing FGDs with a violation of locality constraints.

3.4.1 Resource Limitation Theory

The most prominent processing account of locality constraint is the Resource Limitation Theory (RLT) (Kluender, 1991; Kluender and Kutas, 1993a; Kluender and Kutas, 1993b; Kluender, 1998). This account was built on the WM model presented in the Capacity Constrained Comprehension Theory (Just and Carpenter, 1992). This WM model assumes that verbal working memory has two functions: a computational function and a storage function. Since these two functions draw on a shared pool of resources, the efficiency of one task is negatively affected when demands on the other task increase; thus, storing elements in WM and computing new information requires adequate WM for these tasks to be efficiently executed. Increased demands on WM resources prompt parsers to free up mental resources by either removing stored items or abandoning certain computational processes.

To account for the processing of FGDs, the Resource Limitation Theory states that, upon encountering a filler phrase, parsers keep this phrase active in WM until the integration site is identified. Holding the filler phrase active in WM while searching for the gap site is expected to tax WM resources. Consequently, parsers attempt to resolve this dependency as early as possible. This hypothesis is supported by the findings of experimental investigations, which reveal that processing FGDs is more complex than processing sentences that do not contain FGDs. For example, in a self-paced reading task, Chen et al. (2005) investigated how the processing of sentences created with similar lexical items can be affected by whether or not an FGD is present. These authors manipulated the type of structure between a relative clause, where a

filler phrase needs to be linked with a gap, and a sentential complement, where there is no FGD, as in the following examples:

28. a. The announcement [which the baker from a small bakery in New York City received ____] helped the business of the owner.
- b. The announcement [that the baker from a small bakery in New York City received the award] helped the business of the owner.

Chen et al. (2005) found a slowdown in reading time in (28a) at the critical region 'the baker from a small bakery in New York City' in comparison to (28b), where no *wh*-dependency is established. Similarly, Clifton and Frazier (1989) identified a slowdown in reading time in a self-paced reading task in *wh*-questions compared to their counterparts' yes/ no questions. A set of sample experimental items is presented in (29).

29. a. What did your beautifully dressed niece mutter _ to Willy in the house?
- b. Did your beautifully dressed niece mutter something to Willy in the house?
- c. What did your beautifully dressed niece mutter to Willy about _ in the house?
- d. Did your beautifully dressed niece mutter to Willy about something in the house?

Similarly, Kluender and Kutas (1993a) reported a contrast between *wh*-questions and yes/no questions in the Event-Related Potential (ERP) paradigm. Left Anterior Negativity (LAN) was elicited 300 to 500 ms after a filler phrase was encountered in *wh*-questions (e.g. at the word 'she' in (30a)), compared to a yes/no question where no filler phrase is stored in WM in (30b). LAN was also observed at the integration site in *wh*-questions.

30. a. What has she forgotten that he dragged her to ____ on Christmas Eve?
- b. Has she forgotten that he dragged her to a movie on Christmas Eve?
- (Kluender, 1998)

The contrast between processing sentences with and without FGDs is commonly attributed to the fact that parsing FGDs requires holding a representation of the filler phrase in WM until the integration site is reached.

Furthermore, RLT argues that demands on WM resources are affected by whether or not a clause boundary intervenes between the filler phrase and its associated gap position, as storing a filler phrase while crossing a clause boundary consumes more resources than storing a filler phrase within a clause of the same length. In other words, the process of looking for the gap position while actively maintaining the filler phrase in WM and simultaneously processing intervening clause boundaries is highly resource-consuming.

The increased complexity associated with processing long dependencies in comparison to short dependencies has been widely documented in experimental psycholinguistic research. In a self-paced reading task, Frazier and Clifton (1989) found that dependencies crossing clause boundaries (31b) are associated with a slowdown at the region of the dependency resolution compared to FGDs, where no clause boundary is crossed (31a); this is despite the fact that both conditions were designed to have identical string length.

31. a. What did Katie and Tom mail to New York?
- b. What did Sue think Tom mailed to New York?

Moving on to island effects, RLT argues that processing FGDs requires maintaining the filler phrase in WM throughout the processing of the dependency until it can be reintegrated into the gap position in the embedded clause. Maintaining the filler phrase in WM while processing multiple clause boundaries increases demands on WM resources, resulting in an embedding effect. Island-violating dependencies induce a further source of complexity. Specifically, Kluender (1998) claimed that the inherent semantic complexity or referential properties introduced by the embedded island structures consumes WM resources. Kluender (1998) conducted an acceptability judgement experiment to investigate the effect of clause boundary type on FGD processing. The type of element that introduced the embedded clause boundary was manipulated both in yes/no questions (where there is no FGD) and in *wh*-questions (where an FGD needs to be resolved). See (32).

The semantic and referential properties of the complementisers vary across different types of embedded clauses. For example, the complementiser 'that' in (32a) introduces an embedded non-island structure, which is neutral in terms of its semantic content and merely introduces a proposition; however, the complementiser 'if' in (32b) introduces an adjunct island, which is semantically richer when compared with 'that' and expresses a possible state of affairs. The interrogative pronoun 'what' introduces

a *wh*-island (32c); in contrast to the previous two types of complementiser, this expression has a referent in discourse.

The results of acceptability judgements revealed that the type of element introducing the embedded clause significantly affects acceptability: specifically, 'that' was found to be the most acceptable, followed by 'if' and 'what'. This effect was observed regardless of whether or not an FGD was present. Furthermore, a main effect for structure type was also found, such that *wh*-questions (33) were rated lower than yes/no questions (32), suggesting that FGDs independently increase processing cost.

32. a. Has she forgotten [that he dragged her to a movie on Christmas Eve?] >
b. Has she forgotten [if he dragged her to a movie on Christmas Eve?] >
c. Has she forgotten [who he dragged ___ to a movie on Christmas Eve?] >
33. a. What has she forgotten [that he dragged her to ___ on Christmas Eve?] >
b. What has she forgotten [if he dragged her to ___ on Christmas Eve?] >
c. What has she forgotten [who he dragged ___ to ___ on Christmas Eve?]

(Kluender, 1998)

Hence, resource limitation accounts have argued that maintaining the filler phrase in WM while processing island structures, the semantics of which are inherently complex, overloads the parser's language processing resources; consequently, integrating the filler phrase at the gap position is likely to fail, resulting in perceived unacceptability. In essence, RLT contends that the unacceptability of island-violating sentences arises because they 'involve numerous processing pressures that aggregate to derive processing difficulty above a threshold' (Hofmeister and Sag, 2010, p.366). These pressures are as follows: (i) storing a filler phrase in WM; (ii) processing a clause boundary while holding the filler phrase in WM; and (iii) the inherent semantic complexity of different types of island structures.

3.4.2 Resumption in WM-based accounts of locality constraints

According to processing-based accounts of resumption, the use of RPs is limited to contexts where processing complexity is high. These contexts include deeply embedded dependencies and island-violating dependencies. Experimental investigations in psycholinguistic research support this hypothesis.

Ackerman et al. (2018) used a forced-choice task and a forced-choice sentence completion task to investigate resumption in island and non-island contexts. Experimental sentences included both island and non-island dependencies. Three types of island were investigated in this experiment: Relative Clause Island (34), Adjunct Island (35), and *wh*-Island (36). In the forced-choice task, participants were asked to explicitly compare a minimal pair that differed only in terms of whether a gap or an RP appeared at the dependency tail. In the forced choice sentence completion task, moreover, participants were asked to complete a sentence with a phrase containing either a gap or a RP.

34. a. Island condition

Which woman did Carlos report that the newscaster who exposed threatened the detective's case?

Which woman did Carlos report that the newscaster who exposed her threatened the detective's case?

b. Non-island condition

Which woman did Carlos report that the newscaster who exposed the criminal threatened?

Which woman did Carlos report that the newscaster who exposed the criminal threatened her?

35. a. Island condition

Which woman did Carlos report that, when the newscaster exposed, the criminal threatened the detective's case?

Which woman did Carlos report that, when the newscaster exposed her, the criminal threatened the detective's case?

b. Non-island condition

Which woman did Carlos report that, when the newscaster exposed the detective's case, the criminal threatened?

Which woman did Carlos report that, when the newscaster exposed the detective's case, the criminal threatened her?

36. a. Island condition

Which woman did Carlos question how the newscaster exposed?

Which woman did Carlos question how the newscaster exposed her?

b. Non-island condition

Which woman did Carlos report that the newscaster exposed?

Which woman did Carlos report that the newscaster exposed her?

Results revealed that RPs were significantly preferred to gaps in islands, but dis-preferred in non-island contexts. The facilitation effect of RPs was further found to be significant and robust across the different types of island structures.

Furthermore, Beltrama and Xiang (2016) conducted a comprehensibility judgement task in which participants were asked to rate the comprehensibility of dependency structures. Experimental sentences manipulated islandhood, embedding level (two-level, three-level) and tail of dependency (gap vs RP); see (37). The first experiment was conducted on English sentences.

37. Context sentence:

Have you heard? Yesterday there were riots in the streets. Some people were wounded. Look here, they're talking about it in the paper.

Target sentence:

a. No-Island (2 levels of embedding)

This is the boy that the cop who was leading the operation beat {him / ___ } up.

b. No-Island (3 levels of embedding)

This is the boy that the newspaper reports that the cop who was leading the operation beat {him / ___ } up.

c. Island (2 levels of embedding)

This is the boy that the cop who beat {him / ___ } up was leading the operation

d. Island (3 levels of embedding)

This is the boy that the newspaper reports that the cop who beat {him / ___ } up was leading the operation

(Beltrama and Xiang, 2016, p.12)

Results revealed that RPs are more comprehensible than gaps inside islands under both two-level ($\beta = 0.38$, $p < 0.0001$) and three-level embedding conditions ($\beta = 0.22$, $p < 0.05$). In non-island contexts, on the other hand, gaps are perceived as more comprehensible than RPs both with two-level ($\beta = 0.52$, $p < 0.0001$) and three-level embedding ($\beta = 0.33$, $p < 0.001$). However, gapped dependencies were affected by embedding, such that gaps with two-level embedding were more comprehensible than those with three-level embedding ($\beta = -0.12$, $p < 0.01$). RP dependencies, on the other hand were not affected by level of embedding (all $p > .5$).

The same experiment was conducted for Italian. In non-island contexts, gaps received higher ratings than RPs at two levels of embedding (2-embedding: $\beta = 0.74$, $p < 0.0001$). However, both RPs and gaps were equally comprehensible at three levels of embedding ($\beta = -0.02$, $p > 0.8$), meaning that the comprehensibility rating of RPs increases as the embedding level increases. Inside island structures, moreover, RPs received higher rating than gaps at both two ($\beta = 0.26$, $p < 0.01$) and three levels of embedding ($\beta = 0.35$, $p < 0.001$).

Hofmeister and Norcliffe (2013) investigated the interaction between resumption and embedding. In the long conditions (such as (38a) and (38b)), two clause boundaries separate the filler phrase and its associate verb. In the short conditions (e.g. (38c) and (38d)), only one clause boundary intervenes between the filler and its associate verb. Results revealed that resumption significantly speeds up reading time (RT) at the critical region (the two words following the integration site). However, the faster average RT for resumption is primarily caused by the long condition ($t = -3.85$). Moreover, results revealed that while RT in the gapped condition slows down as the embedding level increases, processing speeds up in RP dependencies with deeper embedding ($t = 2.21$).

38. a. Mary confirmed that there was a prisoner who the prison officials had acknowledged that the guard helped to make a daring escape.
- b. Mary confirmed that there was a prisoner who the prison officials had acknowledged that the guard helped him to make a daring escape.
- c. The prison officials had acknowledged that there was a prisoner that the guard helped to make a daring escape.

- d. The prison officials had acknowledged that there was a prisoner that the guard helped him to make a daring escape.

(Hofmeister and Norcliffe, 2013, p.5)

Hammerly (2019) investigated the extent to which RPs in English facilitate processing in a self-paced reading task in both island and non-island contexts. He manipulated two factors: *islandhood* (non-island/*wh*-island) and *dependency* (gap/RP). See (39), where critical and spillover regions are labelled.

39. a. Non-Island:

Mary trained the spy who Beth announced that the agency_{crit} had recruited {him,}_{spill over} the summer for the program.

- b. *WH*-Island:

Mary trained the spy who Beth announced which agency_{crit} had recruited {him,}_{spill over} the summer for the program.

(Hammerly, 2019, p.4)

Results revealed the main effect of islandhood (Estimate (SE) = -54.96 (25.35), $t = -2.16$, $p = 0.031$): namely, *wh*-islands are read more slowly than non-island conditions. In addition, the results revealed a significant interaction between dependency and islandhood (Estimate (SE) = 143.73 (50.71), $t = 2.84$, $p = 0.005$): RPs slowed down reading compared to gaps in the non-island conditions, but sped up reading in *wh*-islands when compared to gaps.

The longer RTs in the non-island conditions were interpreted as an effect of reanalysis. The parser was searching for a gap, but found a pronoun; hence, a reanalysis process was initiated in which the resumptive pronoun was interpreted as co-referential with the filler phrase. However, as parsers do not expect to find gaps within islands, RPs are initially perceived as intrusive pronouns and are obligatorily interpreted as co-referent with the filler phrase. Hence, Hammerly (2019) hypothesised that the main factor affecting the use of RPs is the 'parser's ability to actively maintain a filler' (p.10); in other words, RPs enable active filling when parsers are unable to keep the filler phrase active in WM.

Chacón (2019) examined the extent to which increasing the load on WM resources improves the acceptability of resumption in FGDs in general. Chacón (2019) hypothesised that, upon encountering a filler phrase, parsers construct a gap-representation and maintain it actively in WM. When the activation level of the

representation is lost due to increased demands on WM resources, parsers generally become less sensitive to unresolved dependencies. Resumptive pronouns facilitate processing as they trigger an anaphoric resolution of the dependency, allowing parsers to retrieve the intended interpretation of the sentence.

Chacón tested this hypothesis in a series of acceptability judgement tasks, finding that the acceptability of RP dependencies significantly improved when the distance between the filler phrase and integration site increased, as well as when participants were engaged in tasks that were demanding in terms of WM resources while simultaneously judging the acceptability of the experimental sentences. Chacón (2019) concluded that when gap representation in WM is lost due to increases in dependency length or increased demands on WM resources, parsers are less likely to observe that the dependency is not resolved. When a resumptive pronoun is encountered, an anaphoric dependency is established between it and the filler phrase. This allows parsers to retrieve the intended interpretation of the dependency. Although Chacón (2019) did not test RPs with island-violating dependencies, the interaction between resumption and island effects naturally follows from this theory, as islands are known to constrain WM resources (Kluender and Kutas, 1993b; Hofmeister and Sag, 2010). Chacón (2018) concludes with the hypothesis that, 'islands that strain working memory resources more are better hosts for resumption' (p.107).

The conclusion that emerges from these experimental studies is that resumption circumvents both embedding and island effects by reducing the processing burden on WM resources. WM-based theories of resumption assume that the facilitation effect of RPs is evident in cases where it is either difficult or impossible to maintain the filler phrase actively in WM. Factors such as the distance between filler phrase and integration site or the crossing of island clause boundaries are known to increase demands on WM. In these cases, RPs are expected to induce a processing advantage, allowing for the creation of a processible and interpretable dependency.

3.5 Evaluating syntactic vs. WM-based accounts of resumption

Cross-linguistic variation poses a major challenge for theories of resumption. The theories that do not distinguish between true and intrusive RPs in terms of their underlying syntactic structures are particularly problematic in this respect; these include syntactic last resort accounts and WM-based accounts.

Despite the strong evidence in support of the WM-based theory of resumption, we present arguments below that this theory cannot account for all the known empirical facts concerning the distributional patterns of RPs. In other words, the view that aims to provide a uniform WM-based account for RPs across languages is challenged by the fact that complementiser type affects whether a dependency ends with a gap or a RP.

Furthermore, WM-based accounts alone cannot explain the distinction between the behaviour of true and intrusive resumptive pronouns. If one attempts to extend this analysis to resumption in true resumption languages, it is unclear why resumption is more facilitative of processing in true resumption languages than in intrusive resumption languages.

Similarly, the last resort theory, which attempts to provide a uniform syntactic account for RPs cross-linguistically, gives no explanation for why RPs are perceived as more acceptable in certain languages than in others.

These facts are easily accounted for with reference to the syntactic theories arguing that an underlying syntactic difference must exist between true and intrusive RP dependencies (i.e. Complementizer type account). McCloskey (2017) argued that the connection between complementiser type and type of dependency tail 'must be represented in the grammar of the language' (p.82). Similarly, Alexopoulou (2010, p.17) claimed that the morphological alternations of complementisers and their connection with the type of dependency structure (i.e. gap or RP dependency) 'mark a distinction between movement and base-generation'. At the same time, the role of WM-related factors should not be ignored. Along these lines, Alexopoulou (2010) provides a structure-based account of the distinction between intrusive RPs and true RPs, and highlights the importance of WM-related factors in accounting for the distributional properties of RPs cross-linguistically. A more detailed discussion of this account is presented below.

3.5.1 Alexopoulou's (2010) theory: True RPs vs intrusive RPs

This theory aims to identify the properties that distinguish the grammar of languages with grammatical resumption from the grammar of those with intrusive resumption. It further hypothesises that grammatical resumption arises through a generalisation of intrusive resumption, which is a strategy for facilitating the processing of complexity that is available to every language. The main features of this theory will be discussed

in 3.5.1.1. We will discuss the validity of the assumptions on which this account is based in 3.5.1.2.

3.5.1.1 Main features of Alexopoulou's (2010) theory

3.5.1.1.1 Grammar of true and intrusive RPs is different

Alexopoulou's (2010) theory relates the distinction between true and intrusive resumption to the morpho-syntactic features of C. Specifically, Alexopoulou (2010) assumes that the distinction between syntactic binding and movement dependencies is key to accounting for the different properties of intrusive and true RPs. Alexopoulou (2010) argued that the correlation between C and resumption is a manifestation of the alternation between movement and binding derivations. Specifically, true RPs appear in dependencies where C obligatorily initiates a binding derivation; intrusive pronouns, on the other hand, appear in dependencies where C obligatorily initiates a movement derivation.

Hence, Alexopoulou (2010) assumed that the source of cross-linguistic variation in the phenomenon of resumption is connected with the morphosyntactic features of C. Whenever C triggers movement, an intrusive pronoun can appear to facilitate processing. True RPs appear in binding dependencies, which are insensitive to islands in the first place.

3.5.1.1.2 Resumption and complexity of processing FGDs

Alexopoulou (2010) outlines a theory of the processing of movement dependencies with intrusive RPs and binding dependencies with true RPs. Her argument is that true RPs arise through a generalisation of intrusive RPs, which have a processing function. The assumptions of the Syntactic Prediction Locality Theory (Gibson, 1998) are adopted for this purpose. This theory states that the complexity associated with processing FGDs is calculated on the basis of two components: (i) *Memory Cost*: the processing cost that arises due to 'the storage of the input that may be used in parsing later parts of an input' (Alexopoulou and Keller, 2007, p.139); and (ii) *Integration Cost*: the processing cost that arises due to integrating an input into an already-built structure. According to Alexopoulou (2010), memory cost depends on distance: i.e. memory cost increases as the distance between the filler phrase and the integration site increases.

However, unlike Gibson's (1998) Syntactic Prediction Locality Theory, which calculates distance in terms of the number of head nouns that intervene between the filler and gap (i.e. distance is measured linearly), Alexopoulou (2010) defined distance in FGDs as the structural distance between the filler and the gap site, calculated in terms of the number of CP nodes between the fillers and gaps. Specifically, Alexopoulou (2010) linked the complexity of movement dependency processing to their successive-cyclic nature. Following Frazier and Clifton (1989), Alexopoulou (2010) assumed that fillers are reactivated/integrated at every intermediate C, due to the presence of intermediate copies or traces of the filler at these intermediate positions in accordance with the successive-cyclic nature of movement dependencies. In essence, this type of analysis attributes memory cost to the successively cyclic nature of movement dependencies. Integration cost, on the other hand, is dependent on both the type of the filler phrase and the (backward) distance between the integration site and its associated filler phrase.

Alexopoulou (2010) used these assumptions to account for the mechanisms by which intrusive RPs facilitate the processing of 'grammatical' movement FGDs. In particular, Alexopoulou (2010) assumed that the parser will anticipate a resolution of the syntactic dependency at a later point in the sentence. Carrying this prediction across several clause boundaries increases memory cost. The presence of an intrusive resumptive pronoun at the integration site cancels the integration cost, since there is no gap to be filled. Once the intrusive pronoun is found at the integration site, the successive-cyclic resolution of the dependency is abandoned; at this point, the parser initiates a reanalysis process in which the dependency between the pronoun and the filler phrase is processed as an anaphoric co-reference dependency.

Alexopoulou (2010) argued that the anaphoric resolution of the dependency is beneficial in terms of processing, as it cancels the distance-based integration cost. However, intrusive resumptive pronouns cannot cancel all processing costs, because until the pronoun is encountered, parsers are carrying a prediction of a gap that increases memory cost. Hence, Alexopoulou (2010) assumed that mixed chains are involved in the derivation of dependencies with intrusive resumption: 'a cyclic derivation associated with the prediction of a trace up to the point of encountering the pronominal when this derivation is abandoned for an anaphoric resolution of the dependency' (p.14).

As for binding dependencies with true resumptive pronouns, Alexopoulou (2010) argued that these dependencies involve non-mixed structures in which all processing,

memory and integration costs, are cancelled. To illustrate, Alexopoulou (2010) assumed that the absence of the *Op* feature in C in binding dependencies means that the operator at spec-C is not processed as filler; in other words, parsers do not hold a prediction that this filler needs to be integrated later into the dependency structure. Thus, there should be no memory cost (i.e. the cost that intrusive resumptive pronouns cannot overcome).

Consequently, the mechanisms that underline the choice between the gap and resumptive strategies (in grammatical contexts) in both true and intrusive resumption languages is determined by processing factors: in short, when a processing complexity arises, the resumptive strategy is chosen over the gap strategy. However, the amount of the processing costs that true and intrusive RPs cancel differs: while true RPs cancel all processing costs, intrusive RPs cancel only the distance-based integration cost.

3.5.1.1.3 Intrusive resumption and islands

As explained above, Alexopoulou (2010) argued that movement is only involved in the derivation of dependencies with intrusive RPs; dependencies with true RPs do not involve movement. Consequently, only intrusive RPs are expected to interact with the complexity of island-violating dependencies. True RPs appear in binding dependencies, which are insensitive to islands by nature. Thus, the following discussion is relevant to intrusive RPs.

The distinction between weak and strong islands is a significant assumption for this theory. Traditionally, the distinction between strong and weak islands is underpinned by the observation that no extraction is permitted in strong islands, while some phrases can be extracted in weak islands (Szabolcsi and Lohndal, 2017). The *wh*-island in (40) is considered a typical example of a weak island, while the Relative Clause Island in (41) is viewed as a typical example of a strong island in English (Chomsky, 1977).

40. ?*Who did Mary wonder whether they will fire t?

41. *Who did John meet the girl who will marry t?

(Alexopoulou and Keller, 2007, p.111)

After observing that crossing non-island and weak-island boundaries induces a slight decrease in acceptability, while crossing a strong island boundary induces a high

decrease in acceptability in Greek, German and English, Alexopoulou and Keller (2007) argued that strong islands involve the violation of grammatical constraints, while weak islands are grammatical (similar to non-island structures) but complex to process.¹ Alexopoulou and Keller (2007) further found that in Greek, German and English, the acceptability of RPs improves as the level of embedding increases in dependencies with embedded *that*-clauses and *wh*-island structures (weak island); see illustrative English examples in (42) and (43) below. On the other hand, RPs are not improved when embedded in relative clause islands (strong islands), as illustrated in the English example in (44) below.

42. Non-island condition (*that*-clause)

- a. Who does Mary claim that we will fire *-/him?* (single)
- b. Who does Jane think that Mary claims that we will fire *-/him?* (double)

43. Weak-island condition (*whether*-clause)

- a. Who does Mary wonder *whether* we will fire *-/him?* (single)
- b. Who does Jane think that Mary wonders *whether* we will fire *-/him?* (double)

44. Strong-island condition (relative clause)

- a. Who does Mary meet the people that will fire *-/him?* (single)
- b. Who does Jane think that Mary meets the people that will fire *-/him?* (double)

Therefore, Alexopoulou (2010) argued that intrusive RPs can only facilitate the processing complexity of non-islands and weak islands and cannot repair the violation of grammatical constraints in strong islands.

It is important to note here that this account assumes that resumption does not save islands per se; rather, islands are sensitive to types of dependency (movement vs binding). When RPs appear in binding dependencies, the acceptability of island-

¹ Alexopoulou and Keller (2007) adopted Szabolcsi and Zwarts's (1993) semantic account of weak islands: weak islands emerge due to an interaction between the scope domain of *whether* with the matrix question's scope domain. The interaction between the two scope domains increase demands on WM resources.

violating dependencies is fully restored. When RPs appear in movement dependencies, they make long-distance dependencies and weak island-violating dependencies interpretable and processible, but not fully acceptable. However, strong islands, as purely syntactic phenomena, are not sensitive to the processing advantage of intrusive RPs.

To summarise, the main features of Alexopoulou's (2010) account are listed below:

- Features of C determine the type of dependency (movement vs binding).
- When C triggers movement:
 - o Intrusive pronouns can ameliorate – but not fully repair – the acceptability of dependencies that cross non-island and weak island clause boundaries.
 - o Intrusive RPs cannot ameliorate island effects resulting from crossing strong island boundaries (because they involve a violation of grammatical constraints on movement, and RPs cannot fix derivational problems).
- When C does not trigger movement, a resumptive pronoun is externally merged and a binding relation established between the pronoun and its related filler phrase to resolve the dependency. Since these dependencies do not involve a violation of grammatical constraints on movement, and further lack the structural complexity of movement dependencies, RP dependencies are expected to be fully acceptable.

3.5.1.2 Assumptions on which this theory is based

Alexopoulou's (2010) theory is based on the following assumptions:

- i. Grammatical constraints have a substantial impact on the theory of island effects (a distinction between movement and binding dependencies must be maintained to account for FGDs' (in)sensitivity to islands).
- ii. Only movement dependencies are subject to the cost of processing complex FGDs (i.e. dependencies crossing non-island and weak island clause boundaries); binding dependencies are not.
- iii. Weak and strong islands do not represent the same phenomenon: the former is a processing phenomenon, while the latter is a syntactic phenomenon.

The validity of these assumptions will be discussed in the following sections.

3.5.1.2.1 Grammatical constraints are substantial in theories of island effects

Although the source of island effects remains controversial, there is strong evidence that grammatical constraints are essential to island effects. The evidence presented below is drawn from studies utilising Sprouse's superadditivity paradigm, and Stowes' filled-gap/gender-mismatch paradigm.

Superadditivity paradigm

In an attempt to tease apart the grammatical and processing factors that contribute to island effects, Sprouse et al. (2013), Kush et al. (2015) and Sprouse et al. (2016) present an experimental 'superadditivity' paradigm that aims to quantify the effects of the extra-grammatical components of island-violating sentences (including processing complexity) and thereby isolate the effect of the grammatical constraints. The extra-grammatical components are as follows: (i) the cost of crossing clause boundaries, and (ii) the cost of processing complex island structures. Two factors are manipulated, each with two levels: structure (island – non-island) and distance between the displaced phrase and its trace (matrix – embedded). This results in four sentence types (or conditions), as shown below:

45. a. Who ___ thinks [that John bought a car]? NON-ISLAND | MATRIX
- b. What do you think [that John bought ___]? NON-ISLAND | EMBEDDED
- c. Who ___ wonders [whether John bought a car]? ISLAND | MATRIX
- d. What do you wonder [whether John bought ___]? ISLAND | EMBEDDED

(Sprouse and Hornstein, 2013, p.314)

The FGD is either short, as in (45a) and (45c), where gaps occur in a matrix clause, or long, as in (45b) and (45d), where gaps occur in an embedded clause. The island structure was found to be either missing, as in (45a) and (45b), existing and not crossed by the *wh*-dependency, as in (45c), or crossed, as in (45d).

In essence, this factorial design allows for investigation of the possibility that the decrease in the acceptability of island-violating dependencies occurs due to processing cost rather than the violation of a grammatical constraint. The effect of

dependency length is calculated as the difference in the mean acceptability ratings of the matrix and embedded FGDs, i.e. (45a) vs (45c).

A further decrease in acceptability might arise from the inherent semantic complexity of embedded island structures (as in (45c)) in comparison to non-island embedded structures (as in (45a)). The structure effect is captured by calculating the difference in mean acceptability ratings between (45a) and (45c).

WM-based accounts assume that island effects will be observed when the independent decreases in acceptability caused by length and island structure are combined within a single sentence, as in (45d). The grammatical component, on the other hand, can be identified as any decrease in acceptability over and above the combined individual decreases caused by the non-syntactic components (i.e. length and type of embedded structure). The island effect is then quantified by subtracting the length and structure effects from the total effect: that is, the island effect is calculated as [(b)-(a)] – [(c)-(a)] – [(d)-(a)]. Any superadditive effect would suggest that the acceptability of this structure is affected not only by these processing factors, but also by grammatical constraints on islands.

Sprouse et al. (2012a) used this design to investigate the source of island effects in English. These authors used a scalar acceptability judgement task and magnitude estimation task to test four island types: adjunct islands, subject islands, complex NP islands, and whether islands. They manipulated both dependency length and the presence of a (crossed/non-crossed) island structure. Sample of experimental sentences is presented in (46).

46. Adjunct Island Example

- a. Who __suspects that the boss left her keys in the car? *NONISLAND/MATRIX*
- b. What do you suspect that the boss left __in the car? *NONISLAND/EMBEDDED*
- c. Who __worries if the boss leaves her keys in the car? *ISLAND/MATRIX*
- d. *What do you worry if the boss leaves __in the car? *ISLAND/EMBEDDED*

Sprouse et al. (2012a) assume that the presence of a single additional processing cost in a sentence is not enough to make parsers perceive it as unacceptable. However, when two are combined in a single sentence, the sum of these costs is

expected to be high enough that parsers will perceive such sentences as not acceptable.

Hence, the sentence in the baseline condition in (46a) is expected to be rated as highly acceptable, since it has neither a long-distance dependency nor an island structure. By contrast, the sentences in (46b) and (46c) are expected to receive a lower rating, as each sentence exhibits a single individual processing cost: long-distance dependency in (46b) and a non-crossed island structure in (46c). The sentence in (46d) contains a long-distance dependency and an island structure simultaneously, resulting in an island-violated structure; this sentence is expected to receive the lowest judgement rating.

We have illustrated above that WM-based accounts predict the decrease in the acceptability judgement rating of island-violating conditions to be equal to the combined decreases in acceptability of a long-distance dependency and an island structure; i.e. WM-based accounts predict that the interaction between dependency length and structure results in a linear-additive effect. For grammatical accounts, on the other hand, the decrease in the acceptability judgement rating of island-violating conditions is greater than the combined decreases in acceptability of a long-distance dependency and an island structure; i.e. grammatical accounts predict that the interaction between dependency length and structure results in a superadditive effect.

However, as noted by Sprouse and Hornstein, 2013, the superadditive effects can also be accounted for under Kluender and Kutas' (1993) resource-limitation theory by assuming that the superadditive effect is 'simply the penalty for exceeding the amount of resources available to the speaker' (Sprouse and Hornstein, 2013, p.29).

Under RLT, the simultaneous processing of both the long *wh*-dependency and the island structure exceeds the parser's limited WM resources and brings forth processing difficulty, leading to a superadditive (rather than linear or additive) decline in the acceptability of island-violated sentences. Under the grammatical account, the superadditive effect reflects the application of a grammatical constraint to the derivation of FGDs, rather than a processing complexity.

As expected under both RLT and grammatical accounts of islands, the findings of Sprouse et al.'s (2012a) acceptability study reveal that *wh*-dependency length interacts with island structure in each of the tested four island types ($p < 0.0001$). This interaction was interpreted as reflecting a superadditive effect.

To further investigate the source of the superadditivity effect, Sprouse et al. (2012a) tested the WM-based theories' hypothesis that the acceptability of island-violating constructions will vary across parsers as a function of individual differences in their WM capacity. Strictly grammatical accounts do not predict any correlation between WM capacity and sensitivity to island effects.

To test the above predictions, the WM capacity of the participants was measured using a serial-recall task in the acceptability judgement experiment and an N-back task in the magnitude estimation experiment. Furthermore, parsers' sensitivity to island effects was calculated using differences-in-differences (DD) scores for each participant. In essence, the DD score measures the effect size of superadditivity. These scores were calculated as the difference between D1 (defined as the difference between the embedded non-island z-score rating (46b) and the embedded island z-score rating (46d)) and D2 (defined as the difference between the matrix non-island z-score rating (46a) and the matrix island z-score rating (46c)).

Calculating the DD score with a sample set of mean ratings:

47. a. $D1 = (\text{Non-Island/Embedded}) - (\text{Island/Embedded})$ (rating: z-score units)

What do you think that John bought ___?	0.5
What do you wonder whether John bought ___? -	-1.5
=	2.0

b. $D2 = (\text{Non-Island/Matrix}) - (\text{Island/Matrix})$

Who ___ thinks that John bought a car?	1.5
Who ___ wonders whether John bought a car? -	0.7
=	0.8

c. $DD = D1 - D2 = 2.0 - 0.8 = 1.2$

(Sprouse et al., 2012, p.92)

The results revealed that there is no relation between the DD scores and WM capacity scores of participants, which is consistent with grammatical-based accounts of islands.

Active-filling paradigm

Frazier and d'Arcais (1989) and Clifton and Frazier (1989) demonstrated the existence of the active-filler strategy in dependency processing. Once a filler phrase

is encountered, parsers actively look for a gap, which is posited at the first potential position allowed by the grammar. This assumes that parsers actively construct FGDs in real time. A wide variety of experimental methods provide supporting evidence for the existence of this mechanism. In a self-based reading task, Stowe (1986) found a slowdown at the direct object 'us' in (48a), which contains an FGD, compared to the sentence in (48b), which has no FGD. Stowe (1986) interpreted this slowdown in reading time as the result of a reanalysis effect that emerges when parsers attempt to posit a gap directly after encountering the verb 'bring', but have to change this initial analysis when encountering the direct object pronoun 'us'. This slowdown in reading time due to this reanalysis process is called the 'filled-gap effect': in essence, a reanalysis process is triggered when parsers are unable to fill the first potential gap site with the filler phrase, as it is already filled with another element. The filled-gap effect indicates that parsers make an early commitment to interpret the filler phrase 'who' as the direct object of the verb 'bring' before it is obvious whether a direct object gap exists. This is always interpreted as reflecting parsers' need to optimise usage of WM resources.

48. a. My brother wanted to know who Ruth would bring us home to ___ at Christmas.
- b. My brother wanted to know if Ruth would bring us home to somebody at Christmas.

In addition to filled-gap effects, supporting evidence for the 'active-filler strategy' has been provided through the use of a plausibility mismatch paradigm, where the plausibility of the filler phrase as an argument of the subcategorizing verb is manipulated. A disruption in the processing of FGDs is obtained at the embedded verb when the filler phrase is an implausible argument for that verb, as in (49). Traxler and Pickering's (1996) eye-tracking experiment found that gaze time is elevated on the verb 'wrote' in (49a) in comparison to (49b). These authors interpreted this disruption in processing as reflecting parsers' attempt to interpret the filler phrase 'the city' as the direct object for 'wrote', showing commitment to the 'active filler strategy', before detecting that this resolution would lead to a semantically implausible interpretation. At this point, parsers initiated a reanalysis process and searched for another potential gap site.

49. a. We like the book that the author wrote unceasingly and with great dedication about ___ while waiting for a contract.

- b. We like the city that the author wrote unceasingly and with great dedication about ____ while waiting for a contract.

The active-filling strategy (Frazier and d'Arcais, 1989; Clifton and Frazier, 1989) is expected to be interrupted inside island structures, as they do not allow gaps. This phenomenon was investigated by Stowe (1986), who examined whether English speakers avoid resolving an FGD inside a complex NP island (bracketed in (50)) in a self-paced reading task.

50. DECLARATIVE

- a. The teacher asked if [the silly story about Greg's older brother] was supposed to mean anything.

WH-EXTRACTION

- b. The teacher asked [what the silly story about Greg's older brother] was supposed to mean ____.

If island constraints on the derivation of English FGDs are applied during real-time processing, parsers are not expected to posit a gap at 'Greg's' in the *wh*-extraction condition, despite the fact that the preposition 'about' has the potential to license a gap. Results indicated that no filled-gap effect was observed at 'Greg's' in the *wh*-question, i.e. that there was no slowdown in RT in comparison to the declarative example. These results suggest that active search for a potential gap site stopped inside islands. Syntactic accounts of island effects take this finding as evidence for the claim that, in real-time processing, parsers respect grammatical knowledge and do not attempt to resolve FGDs if they are not grammatical. Processing-based accounts, on the other hand, assume that parsers suspend the active search for a gap inside island structures, as this would incur a high processing cost (see Phillips, 2013 for further discussion).

Taking these assumptions into account, Keshev and Meltzer-Asscher (2017) examined the extent to which parsers in Hebrew actively resolve filler-gap dependencies inside islands. They hypothesised that, since Hebrew has RPs as a grammatical option to resolve island-violating dependencies, parsers will actively search for the integration site of filler phrases inside Hebrew islands.

First, Keshev and Meltzer-Asscher (2017) used an acceptability judgement task to investigate the extent to which RPs are acceptable inside islands in Hebrew (see (51)). Results revealed that RPs are rated higher than gaps inside Complex Noun

Phrase islands (estimate = -0.46, SE = 0.12, $t = -4$, $p = 0.001$), but show no improvement effects when they occur in Coordinate Structure islands (estimate = -0.09, SE = 0.08, $t = -1.16$, $p = 1$). These results suggest that RPs improve acceptability of Complex Noun Phrase islands, but not Coordinate Structure islands.

51. a. Complex Noun Phrase

Ha-šotrim hekiru et ha-iša še-ha-xašudim
the-cops knew ACC the-woman that-the-suspects
še-takfu ota/- daxafu et ha-melcar
that-attacked her/- pushed ACC the-waiter
be-mis'ada yukratit.
in-restaurant upscale

'The cops knew the woman that the suspects who attacked her pushed the waiter in an upscale restaurant.'

b. Coordinate Structure

Ha-šotrim hekiru et ha-iša še-ha-xašudim
the-cops knew ACC the-woman that-the-suspects
daxafu et ha-melcar ve-takfu ota/_
pushed ACC the-waiter and-attacked her/_
be-mis'ada yukratit.
in-restaurant upscale

'The cops knew the woman that the suspects pushed the waiter and attacked her in an upscale restaurant.'

(Hebrew; Keshev and Meltzer-Asscher, 2017)

Using a filled-gap effect paradigm in a self-paced reading task, Keshev and Meltzer-Asscher (2017) examined the extent to which Hebrew parsers can actively search for an integration site inside the two types of islands. The experimental sentences (see (52) and (53)) include relative clauses and sentential complements with violations of CNP or CS. All relativisation dependencies in relative clauses terminate with resumptive pronouns. There are two potential integration sites between the filler phrase and the integration sites, one inside and the other outside the island structure;

both positions are filled with a noun phrase. In (52) and (53), there is a relativisation dependency between ‘the woman’ and the resumptive pronoun ‘her’. The first region of interest is the filled gap ‘waiter’, and the second is the filled gap ‘the cook’. The control conditions do not include relativisation relations.

Keshev and Meltzer-Asscher (2017) hypothesised that the filled-gap effect should be observed outside islands, and inside CNP islands, but not inside CS islands, since RPs are grammatically licensed only in the former but not the latter. More specifically, parsers are expected to predict an RP upon processing the subcategorising verbs in non-island contexts and inside the CNP island, leading to a filled-gap effect; i.e. an increase in RT upon encountering a NP in these positions.

52. a. Complex Noun Phrase - Relative Clause

Ha-šotrim mekirim et ha-iša še-ha-xašud
the-cops know ACC the-woman that-the-suspect
[še-takaf et ha-melcar] kilel et ha-tabax
[that-attacked ACC the-waiter] cursed ACC the-cook
axrey še-hu daxaf ota.
after that-he pushed her.

‘The cops know the woman who the suspect [who attacked the waiter] cursed the cook after he had pushed her’.

b. Complex Noun Phrase – Sentential Complement

Ha-šotrim imtu et ha-divuax še-ha-xašud
the-cops verified ACC the-report that-the-suspect
[še-takaf et ha-melcar] kilel et ha-tabax axrey še-hu
[that-attacked ACC the-waiter] cursed ACC the-cook after that-he
daxaf et ha-iša .
pushed ACC the-woman.

‘The cops verified the report that the suspect [who attacked the waiter] cursed the cook after he pushed the woman.’

53. a. Coordinate Structure - Relative Clause

Ha-šotrim mekirim et ha-iša še-ha-xašud

the-cops know ACC the-woman that-the-suspect
takaf et ha-melcar [ve-kilel et ha-tabax]
attacked ACC the-waiter and-cursed ACC the-cook]
axrey še-hu daxaf ota.
after that-he pushed her.

‘The cops know the woman who the suspect attacked the waiter and cursed the cook after he had pushed her’.

b. Coordinate Structure-Sentential Complement

Ha-šotrim imtu et ha-divuax še-ha-xašud
the-cops verified ACC the-report that-the-suspect
takaf et ha-melcar [ve-kilel et ha-tabax] axrey
attacked ACC the-waiter [and-cursed ACC the-cook] after
še-hu daxaf et ha-iša.
that-he pushed ACC the-woman

‘The cops verified the report that the suspect attacked the waiter and cursed the cook after he pushed the waman’.

(Hebrew; Keshev and Meltzer-Asscher, 2017)

Keshev and Meltzer-Asscher (2017) found that RTs slow down at the critical words, in non-island contexts, and inside CNP island structures, in relative clauses compared to those in sentential complements. However, no filled-gap effect is observed inside CS islands. These results suggest that the active search for an integration site is grammatically constrained.

From these results, it could be concluded that there is a difference in the underlying grammatical constraint that determines the grammatical status of islands and resumption in Hebrew and English. In other words, when RPs are grammatically licensed inside islands, active search for the integration site is not suspended. These results are problematic for WM-based theories, which assume that active search for an integration site inside islands is suspended, as it would induce an unbearable processing cost.

The contrast between parsers’ expectations for resolving FGDs inside islands in Hebrew and English suggests that the mechanism of resolving island-violating RP

dependencies differs between these two languages. The absence of a filled-gap effect in English suggests that movement is involved in the derivation of these dependencies, and that processing FGDs is suspended inside islands, as there is no means by which the dependency can be resolved grammatically. The presence of RPs enables parsers to retrieve the intended interpretation of the dependency (Hammerly, 2019; Chacón, 2019). On the other hand, the presence of the filled-gap effect in Hebrew suggests that movement is not involved in the derivation of these dependencies. Hence, parsers expect to find a resumptive pronoun inside islands to resolve the dependency grammatically.

Cataphoric dependencies and active search strategy

Cataphoric dependencies are insensitive to island effects, as the dependency between the antecedent and the pronoun is established by binding (Chomsky, 1977); unlike *wh*-dependencies, cataphoric dependencies are licit when the dependency crosses an island boundary (54).

54. His managers revealed that [NP the studio [RC that notified Jeffrey Stewart about the new film]] selected a novel for the script.

(Yoshida et al., 2014)

However, it has been argued that cataphoric dependencies, similarly to *wh*-dependencies, involve an active search strategy: upon encountering a pronoun whose referent is not identified in the preceding discourse, parsers actively search for an antecedent for this pronoun in the following discourse (Kazanina et al., 2007; Fedele and Kaiser, 2014). That is, the cataphoric pronoun triggers an active search for a potential antecedent, in a way similar to that of the filler's active search for a potential integration site in FGDs.

Thus, it is presumably predicted that, under RLT, cataphoric dependencies with the antecedent located inside island structures (as in (54)) should have a similar processing effect on the parser as in island-violating *wh*-dependencies, as crossing an island clause boundary while holding the pronoun active in WM is expected to overload parsers' WM resources. Despite this claim, crossing island boundaries remains grammatical in cataphoric dependencies, but not for *wh*-dependencies. This observation is challenging for processing-based accounts of islands.

However, it has been argued that the parsing process in cataphoric and *wh*-dependencies are distinct, under the assumption that they are processed in different

parsing stages and are therefore differently affected by constraints on WM resources. For instance, Berwick and Weinberg (1986) suggest that constructing a *wh*-dependency occurs during the stage of building the dependency, while relating the pronoun with its antecedent in cataphoric dependencies occurs after the structure of the sentence is fully built. According to this view, *wh*-dependencies and cataphoric dependencies are distinct in terms of processing. This theory can thus explain why *wh*-dependencies and cataphoric dependencies are processed differently in terms of island sensitivity, justifying WM-based accounts of islands.

However, this theory is challenged if cataphoric dependencies are actively interpreted in a way that is similar to the formation of *wh*-dependencies. Yoshida et al. (2014) examined the formation of cataphoric dependencies with the goal of determining whether parsers actively search for an antecedent once a pronoun has been processed. If an active search strategy is involved in the processing of a cataphoric dependency, parsers of cataphoric dependencies will not consider a referential expression to be a potential antecedent of the pronoun if a violation of a grammatical constraint occurs (i.e. when a referential expression is located in a position that involves a violation of binding theory). However, in absence of any violation of binding theory, nothing prevents the parser from interpreting a referential expression to be an antecedent of the cataphoric pronoun, even when it occurs inside island structures.

Yoshida et al. (2014) adopted a gender mismatch paradigm to test these hypotheses. These authors manipulated the gender of the pronoun and the first referential expression. The referential expression can appear inside RC-islands, once with and once without a violation of Principle C. The active search strategy predicts that a gender mismatch effect is observed in RC-islands when there is no violation of Principle C. A sample of experimental sentences is presented in (55).

55. a. No Constraint/Match

His_i managers revealed that the studio that notified **Jeffrey Stewart_i** about the new film selected a novel for the script, but **Annie** did not seem to be interested in this information.

b. No constraint/Mismatch

Her_i managers revealed that the studio that notified **Jeffrey Stewart** about the new film selected a novel for the script, but **Annie_i** did not seem to be interested in this information.

c. Principle C/Match

He_i revealed that the studio that notified **Jeffrey Stewart** about the new film selected a novel for the script, but **Andy_i** did not know which one.

d. Principle C/Mismatch

She_i revealed that the studio that notified **Jeffrey Stewart** about the new film selected a novel for the script, but **Annie_i** did not know which one.

Yoshida et al. (2014) found that RT slows down significantly at the antecedent position 'Jeffrey Stewart' in (55b) compared to (55a), where the coreference relation between the pronoun and the antecedent could be established, as there is no violation for Principle C, although the antecedent is located inside the RC-island mean (RTs: gender match= 438.34 ms, gender mismatch=478.46 ms; $t=3.35$, $p<0.001$). The mismatch effect was not observed when the cataphoric dependency involved a violation of Principle C (gender match = 430.97 ms, gender mismatch = 436.37 ms; $t<1$). These findings confirm that (i) the active search strategy is involved in the formation of a cataphoric dependency, and that (ii) the formation of cataphoric dependencies is not affected by island constraints when there is no violation of the grammatical constraint of Principle C. These results support the argument that islands are not fully reducible to constraints on WM resources, and therefore that grammatical constraints must account for, at least, a proportion of the observed effects.

It is important to highlight the point that, although the experimental literature supports the grammatical theories of islands, we do not assume that island-violating dependencies are not costly in terms of WM resources. In particular, as noted by Alexopoulou (2010), movement dependencies with more level of embeddings across non-island and weak island boundaries are costly in terms of processing. However, we expect that, contrary to Alexopoulou (2010), holding a representation of a filler phrase in WM while processing island clause boundaries will lower the activation level of the filler phrase in WM, regardless of whether the dependency structure is derived by movement or binding relations. Evidence for this claim will be highlighted in the next section.

3.5.1.2.2 Processing movement and binding dependencies

An important assumption of Alexopoulou's (2010) account is that the distance effect is only observed in 'grammatical' movement dependencies (and that intrusive RPs remove the processing cost arising due to distance effect). In particular, Alexopoulou

(2010) assumed that the lack of *Op* feature on C in binding dependencies entails that the element in Spec of C is not a filler (i.e. not an element that needs to be integrated later into the structure), and therefore, that no prediction is associated with it. This means that binding dependencies lack memory cost, as there is no element carried in memory along the dependency structure. Hence, they are not subject to the processing cost related to the structural complexity of movement dependencies; true resumption cancels the movement strategy and the complexity associated with their processing.

However, both experimental findings and empirical facts indicate that binding dependencies, even if they lack the processing cost arising from the structural complexity of movement dependencies, are still sensitive to distance/island effects. That is, the assumption that binding dependencies do not require parsers to maintain filler phrases active in WM throughout the processing of the dependency is not consistent with the findings of psycholinguistic research on FGD processing. Notably, the psycholinguistic investigation of FGD processing does not assume a difference in the real-time processing of movement and binding dependencies; whether a dependency is derived by movement or binding, the filler phrase must be maintained active in WM while clause boundaries are processed until the integration position is reached (Kluender, 1991; Kluender and Kutas, 1993a; Kluender and Kutas, 1993b; Kluender, 1998). In other words, WM-based hypotheses, such as the resource limitation hypothesis, assume that the filler phrase must be maintained active in WM until it is integrated at the gap site, which incurs storage cost regardless of the underlying syntactic structure of the dependency.

Experimental investigation into syntactic dependency processing provides evidence that processing anaphoric dependencies is subject to distance effects. Keshev and Meltzer-Asscher (2019) found a superadditive effect with Hebrew binding dependencies. This means that both island structures and embedded dependencies have an independent processing cost, despite the fact that the dependency is derived by binding relations. Moreover, the observation processing binding dependencies in Hebrew and cataphoric dependencies in English are sensitive to filled-gap effects (Keshev and Meltzer-Asscher, 2017; Yoshida et al., 2014) suggest that filler phrases are maintained active in WM throughout the processing of these binding dependencies. This pattern of effect would not be expected if we were to only adopt Frazier and Clifton's (1989) and Alexopoulou's (2010) structural distance theory.

At the same time, if both binding and movement dependencies are subject to distance effect, as indicated by the above discussion, why are RP structures preferred to gap structures in demanding positions in true resumption languages, despite the fact that both structures are grammatically licensed? This question is better accommodated by Alexopoulou's (2010) structural distance theory, which attributes the cost of processing FGDs to the successive-cyclic nature of movement dependencies (i.e. the requirement to integrate filler phrases at intervening CPs); intrusive RPs can partially circumvent these effects. Binding dependencies, on the other hand, are not derived successive-cyclically, but are rather established in one single step (Boxell, 2012); hence, although they are still subject to general memory costs of processing FGDs, they are not subject to the cost of integrating fillers at intervening CPs. This processing account can thus explain why binding dependencies with true RPs are preferred to movement dependencies with gaps in positions that induce high processing costs in true resumption languages (Ariel, 1999; McCloskey, 2017).

3.5.1.2.3 Weak vs strong islands and intrusive RPs

We illustrated above that the distinction between strong and weak islands is an important assumption for this theory. A common interpretation regarding the nature of this distinction is that strong islands represent a syntactic phenomenon, while weak islands represent a non-syntactic (processing) phenomenon. This is the view adopted by Alexopoulou (2010).

Alexopoulou and Keller (2007) argued that, based on experimental findings, intrusive RPs can aid in alleviating the complexity of processing movement dependencies crossing non-island and weak island clause boundaries (grammatical but structurally complex dependencies), but cannot repair the violation of grammatical constraints in movement dependencies crossing strong islands (ungrammatical structures).

The same pattern of interaction between resumption and strong/weak islands has been reported in other languages. For instance, Sterian (2016) provides examples from some Arabic varieties, where resumption improves the acceptability of weak islands (like *wh*-islands), but does not do so for strong islands (for instance, Complex-NP/Relative clause islands). The following example from Lebanese Arabic shows that the resumptive pronoun *a* ('her') improves the acceptability of *wh*-islands.

56. Wh-island

- a. Sməʕt ʔənnu Naadya, byaʕrfo ʔ ayya walad

heard.1S that Nadia know.3P which boy
 šeef=a.

saw.3S=her

'I heard that Nadia, they know which boy saw her.'

b. *Sməʕt ʔənnu Naadya, byaʕrfo ʔayya walad

heard.1S that Nadia know.3P which boy

šeef ____

saw.3S ____

'I heard that Nadia, they know which boy saw ____.'

(Lebanese Arabic; Aoun and Benmamoun, 1998)

By contrast, the example below is an island-violating dependency, where the dependency crosses a Complex Noun Phrase island boundary, but the resumptive pronoun does not improve its acceptability.

57. COMPLEX NP-ISLAND

*telmiz-a l-kesleen ma badkun txabbro wala

student=her the=bad Neg want.2PI inform.2PI no

mʕallme ʕan l-bent yalli seeʕadit-o b-l-faḥiṣ

teacher.F about the=girl who helped.3FS=him in-the-exam

'Her bad student, you didn't want to inform any teacher about the girl who helped him in the exam.'

(Lebanese Arabic; Aoun et al., 2001)

The same phenomenon is observed in Iraqi Arabic. In (58), a RP saves an island-violating dependency that crosses a *wh*-island boundary.

58. WH-ISLAND

a. ya: ʃu:ra li-bni=ha Samer ysʔal iða kull

which picture.F of-son=her Samer ask.3MS if every

mraya šagagat=ha

woman tore.3FS=it

'Which photo of her son did Samer wonder if every woman tore [it]?

- b. * a. ya: řu:ra.F li-bni=ha Samer ysʔal iða
 which picture.F of-son=her Samer ask.3MS if
 kull mraya řagagat=ha
 every woman tore.3FS=_

'Which photo of her son did Samer wonder if every woman tore _ ?'

(Iraqi Arabic; Sterian, 2016)

On the other hand, the example in (59) below is an island-violated dependency, where the dependency crosses a Complex Noun Phrase island boundary, but the resumptive pronoun does not improve its acceptability.

59. COMPLEX-NP ISLAND

- a. *telmið=ha iš-řa:řer ma triddu:n txabbru:n
 student=her the-smart Neg want.2PI inform.2PI
 wala: muřalma řan il-bnaya illy sařadat=ah
 no teacher.F about the-girl who helped.3FS=him
 bi-l-imtiřa:n
 at-the-exam

'Her smart student, you didn't want to inform any teacher about the girl who helped [him] in the exam.'

- b. *telmið=ha iš-řa:řer ma triddu:n txabbru:n
 student=her the-smart Neg want.2PI inform.2PI
 wala: muřalma řan il-bnaya illy sařadat____
 no teacher.F about the-girl who helped.3FS____
 bi-l-imtiřa:n
 at-the-exam

'*Her smart student, you didn't want to inform any teacher about the girl who helped ____ in the exam.'

(Iraqi Arabic; Sterian, 2016)

As explained above, to account for these observations, Alexopoulou (2010) argued that intrusive RPs circumvent the cost of processing 'grammatical' movement dependencies (i.e. dependencies crossing weak and non-island clause boundaries). Strong islands cannot benefit from the processing advantage of intrusive resumption because they involve a violation of grammatical constraint.

Recently, Chaves and Putnam (2020) questioned the nature of the distinction between weak and strong islands. In brief, these authors provided evidence for the variability in the strength of adjunct islands, which have traditionally been viewed as strong islands. In an acceptability judgement study, Chaves and Putnam (2020) found that the overall acceptability of adjunct island violation in English is as low as the overall acceptability of ungrammatical fillers, and lower than the overall acceptability of grammatical fillers (adjunct islands: 2.45, SD = 1.26; grammatical fillers: 4.16, SD = 1.05; ungrammatical fillers: 2.19, SD = 1.11). At the same time, the acceptability of 'conditional' adjunct islands was found to be less degraded than the acceptability of 'temporal' and 'causal' adjunct islands (conditional: Estimate = 0.007, $t = 2.77$, $p < 0.005$; temporal: Estimate = 0.01, $t = 3.78$, $p < 0.0001$; causal: Estimate = 0.0012, $t = 4.7$, $p < 0.0001$).

Interestingly, Chaves and Putnam (2020) found that the acceptability of adjunct island violations is satiated over repeated exposure (Estimate = 0.008, $t = 6.07$, $p < 0.001$). After nine exposures, adjunct island violations became significantly more acceptable than ungrammatical controls. Conditional items were no longer statistically different from grammatical controls after seventeen exposures. Causal and temporal items required more exposures.

These findings are challenging for the claim that strong islands are purely syntactic phenomena. Consequently, Chaves and Putnam (2020) proposed an eclectic view of islands: in short, they argued that different types of islands arise due to a combination of cognitive factors, not merely syntactic factors, and that the strength of different types of islands is gradient, but not categorical, in nature.

Adopting Chaves and Putnam's (2020) eclectic view of islands, the processing advantage of intrusive RPs is expected to be observed in FGDs crossing different types of islands (i.e. not merely weak islands). Indeed, Ackerman et al.'s (2018) experimental study, introduced in detail in section 3.4.2, provided evidence that intrusive RPs improve the acceptability of strong and weak islands in English. Such findings would not be expected if we were only to adopt Alexopoulou's (2010) account, which treats strong and weak islands as distinct phenomena and attributes the

processing advantage of intrusive RPs to reducing the processing cost associated with the structural complexity of ‘grammatical’ movement dependencies.

Furthermore, Chaves and Putnam's (2020) eclectic account of islandhood would expect to find gradience in the sensitivity of islands (which differ in the level of their strength) to the processing advantage of intrusive RPs. In other words, weaker islands are expected to benefit more from the processing advantage of intrusive RPs than stronger islands. Indeed, this prediction is confirmed in Tucker et al.'s (2019) experimental investigation of RPs in MSA.

Tucker et al. (2019) utilised the superadditivity paradigm to test the extent to which resumption can remove the superadditive effect of different types of island structures on data from MSA (Tucker et al., 2019). In other words, the aim was to test the prediction that the superadditive effect, which indicates the existence of island effects, would be changed into a linearly additive pattern through the occurrence of resumptive pronouns, rather than gaps, as the tail of the dependency. Using the factorial design, the experimental sentences manipulate dependency distance (long vs short) and islandhood (island vs non-island) and resumption (gap vs RP). Different types of islands are tested: Adjunct island, Whether-Island and Complex NP Island. The resumption variable is only manipulated inside the Long dependency conditions, as manipulating it inside the Short conditions would result in questions involving a *wh*-item being immediately followed by a resumptive pronoun (e.g. ‘*Who he wonders if Shawn saw Gus?’) (Tucker et al., 2019). See (60).

60. a. Short, No Island, No Resumption

man jaʕtaqidu ʔanna maħmood saraqa ʔal-miħfaza?

who thinks.3MS COMP Mahmoud stole.3MS the-wallet

‘Who thinks that Mahmoud stole the wallet?’

b. Long, No Island, No Resumption/Resumption

ʔayya maħfazah jaʕtaqidu ʔaš-šarṭii ʔanna maħmood

Which wallet thinks.3MS the-policeman COMP Mahmoud

saraqa ___/hu?

stole ___/it?

‘Which wallet does the policeman think that Mahmoud stole?’

c. Short, No Island, No Resumption

man jatasaaʔalu maaʔiðaa kaana maḥmood
who wonders whether had.3MS Mahmoud
saraqa ___/ ʔal-miḥfaʔa?
stole.3MS the-wallet

‘Who wonders whether Mahmoud stole the wallet?’

d. Long, Island, No Resumption/Resumption

ʔayya maḥfazah yatasaaʔalu ʔaš-šarṭii maa ʔiðaa
Which wallet wonders the-policeman whether
kaana maḥmood saraqa ___/hu?
had Mahmoud stole _/it?

‘Which wallet does the policeman wonder whether Mahmoud stole?’

(MSA; Tucker et al., 2019)

To identify the superadditive effect of each island structure, along with the resumption effect on the superadditivity for a given island, Differences-in-Differences (DD) scores were calculated. Since the experimental design was a defective $2 \times 2 + 2$ design, the authors computed two DD scores. The first one, calculated as the difference between D1 (defined as the difference between the long-gapped conditions, with and without island) and D2 (defined as the difference between the short-gapped conditions, with and without island), aimed to estimate the superadditive component of island structures in gap dependencies. The second one was designed to find out whether the superadditive component of island structures is present in long-distance dependencies with RPs. To achieve this, the score was calculated in the same way as the first one, except that D1 was calculated by taking the difference between the long-distance, gapped conditions without island, and the long-distance, RP-conditions with island. To estimate the ameliorative effect of resumption on the superadditivity of island structures, a comparison was drawn between the two DD scores.

The results revealed that both the presence and the magnitude of the superadditive effect vary across the different island structures. The ameliorative effect is found to be diverse across the different types of islands; while it does not emerge when the resumptive pronoun occurs inside Complex Noun Phrase Islands (0.39 with a gap versus 0.40 with a pronoun), it is found to be clear in Adjunct Islands (no resumption

at 0.78 versus resumption at 0.20), and less clear in Whether Islands (no resumption at 0.20 versus resumption at 0.19).

These findings can be accounted for by adopting Chaves and Putnam's (2020) eclectic view of islands, which argues that the strength of island effects is gradient in nature, suggesting that weaker islands are more sensitive than stronger islands to the processing advantage of non-syntactic factors.

3.5.1.3 Interim conclusion

To summarise, in this section, we introduced Alexopoulou's (2010) theory and highlighted its importance to accounting for the distinction between the grammar of true RPs and that of intrusive RPs. We further assume, in line with both Alexopoulou's (2010) account and the Resource Limitation Theory (Kluender, 1991; Kluender and Kutas, 1993a; Kluender and Kutas, 1993b), that while both movement and binding dependencies are subject to an embedding effect that affects the level of activation of a filler phrase in WM, only movement dependencies are subject to the processing cost caused by the need to integrate fillers at intermediate CP. Hence, we speculate that 'grammatical' movement dependencies are costly in terms of processing compared to binding dependencies due to the processing complexity incurred by integrating filler phrases at intermediate CP. We speculate that this might explain why binding dependencies with true RPs are preferred to movement dependencies with gaps in complex-to-process structures in true resumption languages.

Intrusive RPs, on the other hand, help to alleviate the costs of processing 'grammatical' movement dependencies (i.e. the distance-based integration cost) (Alexopoulou, 2010), as well as the processing costs related to the constraints on WM resources (Erteschik-Shir, 1992; Hammerly, 2019; Chacón, 2019).

We further show that the empirical facts and experimental findings call into question the categorical distinction between strong islands (as syntactic phenomena) and weak islands (as processing phenomena), and support a view that this distinction is gradient rather than categorical in nature, as well as that cognitive factors, in addition to grammatical factors, contribute to island effects (Chaves and Putnam, 2020). We assume that this eclectic view of the nature of islands holds that weaker islands are more sensitive to non-syntactic factors, such as the processing advantage of intrusive RPs, than stronger islands.

3.6 Type of dependency structure and resumption

The acceptability of RPs is affected by the type of dependency structure in which they appear. Cross-linguistically, RPs are more common in relative clauses, but restricted in *wh*-questions, where their availability is subject to the discourse properties of filler phrases (Alexopoulou, 2010). For instance, in Lebanese Arabic, resumptive pronouns are obligatory in relative clauses (61). In *wh*-questions, RPs are either optional, as in (62) or prohibited, as illustrated in (63).

61. "am fatti' ʕa kteeb (*yalli) Dayyaʕt-o l-yom
 Asp look.1S for book that lost.1S-it today
 'I am looking for a book that I lost today.'

(Lebanese Arabic; Aoun et al., 2001)

62. a. ayya mmasil šeft-___ b-l-maʕʕam?
 which actor saw.2MS___ in-the-restaurant
 'Which actor did you see in the restaurant?'
- b. ayya mmasil šeft=o b-l-maʕʕam?
 which actor saw.2MS=him in-the-restaurant
 'Which actor did you see in the restaurant?'

(Jordanian Arabic; Abdel Razaq, 2011)

63. *šu štarayt=i ?
 what bought.2SF=it
 'What did you buy?'

(Jordanian Arabic; Abdel Razaq, 2011)

Furthermore, while Hebrew allows resumption with relative clauses, it does not allow resumption in *wh*-questions, as can be seen in (64).

64. a. m mi nifgaSta
 with who you met?
 b. *mi nifgaSta ito
 who you-met with-him?
 'Who did you meet with?'

(Hebrew; Sharvit, 1999, p.591)

However, Sharvit also noted that some varieties of colloquial Hebrew allow resumption with discourse-linked fillers, as in (65):

65. e student nifgaSta ito
which student you-met with-him
'Which student did you meet with?'

(Hebrew; Sharvit, 1999, p.591)

In Somali, on the other hand, RPs are obligatory in *wh*-questions (regardless of the discourse properties of *wh*-fillers) (Hedding, 2014). The following examples illustrate this point.

66. a. **[Nin-kee]_F** b=aad sheegtay in=*(**uu**)
man-which FOC=2SG report.PST.2SG C=3SG.M
ku caayey?
2SG insult.PST.3SG

'Which man did you say insulted you?'

(Somali; Hedding, 2014)

- b. **[Kum]_F=ay** u maleynaysaa in=*(**uu**) Amina
who.FOC=she to think that=3SG.M Amina
arkay?
see.PST.3SG.M

'Who does she think saw Amina?'

(Somali; Saeed, 1984, p.144)

- c. **[Max]_F=aad** doonaysaa in=*(**ay**)
what.FOC=2SG want.PROG.2SG C=3PL
dhacaan?
happen.PRES.3PL

'What do you want to happen?'

(Somali; Saeed, 1984, p.144)

This section illustrates that the type of dependency structure affects the availability of RPs in a given language.

3.7 Summary of our initial assumptions

In summary, following a detailed review of the descriptive and quantitative findings of resumption across languages, we will adopt the following as our initial hypotheses regarding the grammatical and processing constraints governing RP distribution in FGDs:

- Sensitivity to islands is a diagnostic for syntactic movement.
- Features of C affect the type of dependency; this implies that resumption behaviour differs across different types of dependency structures.
- True RPs appear in binding dependencies and fully restore their acceptability.
- Intrusive RPs appear in movement dependency and partially improve acceptability.
- The extent to which intrusive RPs can ameliorate islands varies across types of island: weaker islands are more sensitive to the processing advantage of resumption than stronger islands.

3.8 Research questions

The first part of this dissertation aims to address the following research questions:

1. What are the types of RPs in Baha Arabic?
2. What is the role of RPs in processing across different types of dependency structures in Baha Arabic?

The first experimental study in Chapter 4 is designed to address the first research question. The second research question is addressed in the online studies presented in Chapter 5.

The second part of this dissertation is concerned with investigating the discourse properties of RPs inside islands in Baha Arabic. The related literature and theoretical assumptions relevant to this topic will be presented in Chapter 6.

3.9 Methodological concerns

3.9.1 Formal vs informal acceptability judgements

The judgement of acceptability is the main empirical foundation on which syntactic theories are based (Schütze, 1996). There are currently two methods used to collect acceptability judgements: the traditional informal method, which is dependent on introspective judgements, and the formal method, which employs experimental data collection techniques.

Informal acceptability judgements are usually based on the researcher's introspective judgements, along with those of a limited number of speakers. Phillips (2009) claimed that widely accepted generalisations in formal syntax are not derived based on misleading premises; according to Phillips (2009), 'empirical claims... undergo extensive vetting before they attain the status of "widely accepted generalisation"'.

Despite the fact that the advantages of informal acceptability judgements are evident, many linguists have called for the use of more systematic methods in collecting acceptability judgements to supplement the informal methods used in syntactic literature (Coward, 1997; Sprouse, 2007). Experimental acceptability judgement studies are conducted using a large number of informants, who are asked to rate the acceptability of a number of syntactically identical sentences. Proponents of experimental methods have argued that larger sample sets (of both participants and linguistic items) can mitigate the variability and increase the reliability of data derived on the basis of small sample sizes (Schütze, 1996; Coward, 1997; Sprouse, 2007).

According to Myers (2009), the quantitative analysis of multiple data points undoubtedly offers a more accurate and reliable interpretation than the qualitative analysis of a few individual data points. Moreover, the use of statistical methods enables researchers to evaluate both the trustworthiness and reliability of the obtained results through the calculation of statistical significance (Myers, 2009). Hence, one important dimension for favouring formal methods to informal methods is the availability of the use of statistical methods that test reliability.

Furthermore, as noted by Francom (2009), while informal acceptability judgements are sufficient in uncontroversial cases, such in (67) and (68), these methods are less capable of reliably evaluating gradient cases of acceptability, such as that in (69).

67. Well-formed

Agatha Christie has written many books. b. I don't like detective stories.

68. Ill-formed

*Agatha Christie many books written has. b. *I detective stories like.

(Francom, 2009, cited in Haegeman, 1994)

69. a. ??Which car did John ask how Mary fixed?

b. ?Who did John ask which car fixed?

(Francom, 2009)

Moreover, accommodation effects, as the one observed with the violation of adjunct islands in English (Chaves and Putnam, 2020), represent a much bigger problem for informal methods for collecting acceptability judgments.

Proponents of formal methods do not in fact argue against the importance of informal methods in establishing the empirical facts of syntactic theories. However, they do contend that experimental methods allow for more reliable and fine-grained observations to emerge (Keller, 2000; Sorace and Keller, 2005).

3.9.2 Gradience in acceptability

It is well-known that the relation between grammaticality and acceptability is not explicit (Phillips et al., 2019). More specifically, while acceptability judgements are known to be graded in nature, linguistic theories created by syntacticians tend to be categorical in nature (Keller, 2000; Sorace and Keller, 2005). For their part, grammatical theories create a binary distinction between well-formed structures, which are derived by the underlying grammatical competence system, and ill-formed structures, which cannot be generated by the competence system. A linking hypothesis between acceptability and grammaticality assumes that if a sentence sounds acceptable, it represents a well-formed structure; ill-formedness, on the other hand, is manifested in the form of degraded acceptability.

The widely accepted assumption adopted by most linguists is that the gradience of acceptability reflects interaction between grammatical knowledge and external (non-grammatical) factors, such as semantic anomalies, discourse- or processing-related factors; non-grammatical factors may lead to a mild decrease/increase in acceptability, while violating grammatical constraints is expected to lead to a strong degradation of acceptability.

This account of gradience is primarily connected with Chomskyan notions of competence and performance. In particular, according to Chomsky, formal syntactic theories are centred on the 'ideal speaker-listener' who 'is unaffected by such grammatically irrelevant conditions as memory limitations, distractions, shifts of attention and interest, and errors (random or characteristic) in applying his knowledge of this language in actual performance' (1965, p.3). That is, syntacticians are mainly interested in constructing models that reflect competence alone, with categorical distinctions between well-formed (grammatical) and ill-formed (non-grammatical) structures. However, as noted by Chomsky himself, the effect of performance pressure on acceptability is not avoidable.

Adopting this perspective, intrusive resumption, as a performance phenomenon, might slightly improve the acceptability of sentences with island violation. True resumption, on the other hand, as an integrated part of the grammar of a given language, is predicted to alter the pattern of acceptability: when the presence of such structures is grammatically required, they are fully acceptable, while their absence in these cases will result in highly degraded acceptability.

It is important to highlight here that since acceptability judgements reflect the sum of ratings a condition receives from all participants, a gradient of acceptability might actually reflect significant inter-individual variation rather than a consistent intermediate rating. For instance, Kush et al. (2017) suggest that some cases of gradient patterns of acceptability, as in the case of islands in Mainland Scandinavian languages, might actually arise from significant inter-individual variation in the sensitivity to *wh*-island effects and not from a stable intermediate rating. Kush et al. (2017) found that acceptability of *wh*-islands in these languages is within 'acceptability range'. Hence, Kush et al. (2017) claim that *wh*-islands in these languages involve the violation of a semantic constraint, but not a syntactic constraint (as in English). In particular, these authors suggest that inter-individual variation in the sensitivity to *wh*-island effects results from a variation in the interpretation given to the phrase 'whether' ('om', in Norwegian), which can either act as an operator or a non-operator. This account accordingly assumes that *wh*-islands in Mainland Scandinavian languages are grammatically licensed, but that the illusiveness of ungrammaticality is due to the different interpretation given to 'whether'.

However, it is important to note that Kush et al.'s (2017) findings contrast with early work on gradient acceptability, as in (Cowart, 1997), where the judgements with intermediate values are consistent across and within subjects. To tease apart inter-

individual variation from ‘genuine’ gradience, linguists depend on the distribution of the acceptability data both within and between subjects.

3.10 Conclusion

In this chapter, we have reviewed the syntactic and psycholinguistic accounts of resumption. We reasoned that, following Alexopoulou (2010), a distinction should be maintained between true and intrusive resumption dependencies in terms of their underlying syntactic structures: while successive-cyclic movement is involved in the derivation of intrusive RP dependencies, binding relations are involved in the derivation of true RP dependencies. Furthermore, the assumptions of WM-based theories are essential to accounting for the distributional patterns of RPs cross-linguistically. Based on this review, we identified the following factors that might determine the type of RPs, as well as their distributional patterns, in a given language: (i) type of complementiser; (ii) dependency distance; (iii) crossing island boundaries; (iv) type of island structure; and (vi) type of dependency structure. In the next chapter, we will investigate how these factors contribute to determining the type of RPs used in Baha Arabic, as well as their distributional patterns.

Chapter 4 Acceptability of RPs in Baha Arabic *Wh*-dependencies

4.1 Introduction

This chapter presents an acceptability judgement study of resumption in Baha Arabic. It has one main objective, which is to systematically investigate the use of gaps vs RPs in *wh*-dependencies in Baha Arabic. For this purpose, we will investigate the differences in behaviour of RPs across different types of *wh*-dependencies. More specifically, we aim to investigate how dependency length and the violation of island constraints affect the distributional pattern and the magnitude of acceptability of resumption across different types of *wh*-dependencies. We will further investigate the extent to which the acceptability of RPs correlates negatively with participants' Working Memory (WM) capacity.

In what follows, we introduce how resumption is expected to interact with islands, dependency length and WM capacity in different types of *wh*-dependencies in Baha Arabic, as informed by the literature review. The remainder of this chapter is organised as follows. Section 4.2 briefly reviews the factors affecting the distributional properties of RPs across different *wh*-dependencies in Baha Arabic. Section 4.3 presents the predictions made regarding the behaviour of RPs across different types of dependency structures in Baha Arabic. Section 4.4.1 presents the methods of the current experiment, while Section 4.4.2 presents the results and a discussion of the data. Sections 4.4.4, 4.4.5 and 4.4.6 summarise these findings, discuss any potential limitations and outline the outstanding questions. Finally, Section 4.5 concludes the chapter.

4.2 Background

4.2.1 Resumption and islands

As discussed above, the theoretical literature contends that resumptive pronouns cross-linguistically have an amelioration effect on the acceptability of island-violating dependencies (Ross, 1967; Erteschik-Shir, 1992; Shlonsky, 1992; Rouveret, 2011; Asudeh, 2012). However, there is no consensus on the magnitude of this amelioration effect; in other words, the exact status of the acceptability of these island-violating dependencies with RPs is still not clearly defined. For instance, in English, some use notation conventions for acceptability that indicate full acceptability of RP

dependencies, as in (1). Others use the ? diacritic that indicates an intermediate status of acceptability, as in (2). In true resumption languages, researchers are in agreement that RPs fully repair dependencies with island violations, as in (3). However, this prediction was not replicated in experimental studies (Farby et al., 2010; Meltzer-Asscher et al., 2015; Tucker et al., 2019). We assume that the magnitude of the amelioration is an indicator of the nature of the effect. In cases where RPs are syntactically licensed, their presence will result in highly acceptable island-violating dependencies. On the other hand, when RPs are required for processing function, their presence will result in only mild improvement in the acceptability of island-violating dependencies.

1. a. *I just saw a girl who Long John's claim that _____ was a Venusian made all the headlines.
b. I just saw a girl who Long John's claim that she was a Venusian made all the headlines.
(Ross, 1986, p.260)
2. a. *These are the things that we do not know what __ are.
b. ?These are the things that we do not know what they are.
(Chacón, 2015, p.92)
3. a. ya: Su:ra.F li-bni=ha Samer ysʔal iða
which picture.F of-son=her Samer ask.3MS if
kull mraya ʃagagat=ha
every woman tore.3FS=3F
'Which photo of her son did Samer wonder if every woman tore [it] ?'
b. * ya: Su:ra.F li-bni=ha Samer ysʔal iða
which picture.F of-son=her Samer ask.3MS if
kull mraya ʃagagat__
every woman tore__
'Which photo of her son did Samer wonder if every woman tore __ ?'
(Iraqi Arabic; Sterian, 2016)

In the present study, we attempt to provide evidence concerning three major questions in the literature on the relationship between resumption and island amelioration, as follows: (1) Do RPs in true resumption languages (such as Baha Arabic) exhibit a different pattern of amelioration effects on islands, compared to languages with intrusive RPs (such as English)? (2) Is the magnitude of the amelioration effect of resumption affected by island type? (3) Is the magnitude of the amelioration effect of resumption affected by dependency structure type?

The first and second questions have already been addressed experimentally in Modern Standard Arabic (henceforth, MSA) (Tucker et al., 2019). The results of Tucker et al.'s (2019) study indicate that while the magnitude of the amelioration effect of resumption varies across different types of islands, this amelioration effect is described as 'negligible'. However, as stated in section 2.4, one shortcoming of this study is that the participants' proficiency in MSA might be not sufficient to allow them to make such acceptability judgements; it is frequently claimed that there are no native speakers of MSA (Kaye, 1970; Maamouri, 1998), since the use of MSA is restricted to formal occasions. Tucker et al. (2019) speculate that the lack of formal assessment of MSA proficiency in their study might have affected the results. Furthermore, Tucker et al.'s study only investigated *wh*-questions. As noted earlier, the use of RPs cross-linguistically is common in relative clauses, but restricted in *wh*-questions (Boeckx, 2003; Alexopoulou, 2010).

In the present study, we will address these questions through a study of native speakers of Baha Arabic, a spoken variety of Arabic. We will further extend the range of structures to include relative clauses and two types of *wh*-questions: specifically, cleft-*wh* questions and bare *wh*-questions. The structural properties of these types of *wh*-dependencies are illustrated in (4) below. In essence, the theoretical literature on resumption in different varieties of Arabic contends that the grammatical status of RPs differs across these different *wh*-dependencies, as shown in (4). RPs are predicted to be fully acceptable inside islands in relative clauses (4a) and cleft-*wh*-questions (4b), but unacceptable in bare *wh*-questions (4c).

4. a. Relative clause:
Head noun phrases + Complementiser *illi*RP/*Gap
- b. Clefted *wh*-questions
Wh-phrase+(NP)+ Complementiser *illi*RP/*Gap?
- c. Bare *wh*-question

Wh-phrase Gap/*RP?

These different patterns of acceptability are interpreted as reflecting different underlying syntactic structures across the three types of *wh*-dependencies. In particular, the sensitivity of bare *wh*-questions to islands, in both the absence and presence of RPs, was taken as evidence that derivation of these structures involves movement (Aoun et al., 2009). On the other hand, insensitivity to islands in relative clauses and clefted *wh*-questions with RPs was taken as evidence that deriving these structures does not involve movement (Aoun et al., 2009).

These facts are consistent with Alexopoulou's (2010) hypothesis that resumption, *per se*, does not repair islands; in other words, what Alexopoulou (2010) deems to matter here are the underlying syntactic structures of a *wh*-dependency. Islands are fully acceptable if RPs appear in binding dependencies. When RPs appear in movement dependencies, they partially ameliorate islands. However, Arabic linguists do not use diacritics that illustrate the amelioration effect of RPs in bare *wh*-questions; this suggests that Arabic linguists assume RPs to be ungrammatical and likely to have no processing effect in Arabic bare *wh*-questions.

Alexopoulou's (2010) hypothesis is further supported by the fact that relative clauses and cleft *wh*-questions are obligatorily introduced with the complementiser *illi*, while bare *wh*-questions lack this complementiser. In particular, Alexopoulou (2010) linked the type of complementiser to the type of underlying syntactic structure of *wh*-dependencies. Of particular interest to us here is the Arabic complementiser *illi*. Alexopoulou (2010) claimed that the Arabic complementiser *illi* always lacks the *Op* feature that triggers movement. In this case, the operator is merged in its surface position and a binding relation is established between it and a pronominal at LF. This explains the obligatoriness of RPs in cleft *wh*-questions and relative clauses.

These claims are based on the theoretical syntactic literature. In the present study, our goal is to investigate whether relative clauses and cleft *wh*-questions with resumptive pronouns are sensitive to islands. We further test the extent to which RPs in movement dependencies, as bare *wh*-questions, ameliorate island effects.

4.2.2 Resumption and island type

As explained in the previous chapter, the amelioration effect of intrusive RPs is affected by the strength of the type of island structure in which they occur. Specifically, the processing advantage of RPs is expected to be more evident in weaker islands

than in stronger islands. We will address this issue in the present study by investigating the different extents to which different types of islands affect the amelioration effect of RPs in Baha Arabic. The islands included in this research are adjunct and relative clause islands, both of which are traditionally considered strong islands (Szabolcsi and Lohndal, 2017). However, the status of adjunct islands as strong islands has also been questioned (Cinque, 1990; Truswell, 2007; Farra, 2019; Chaves and Putnam, 2020). For instance, Chaves and Putnam (2020) found that adjunct islands, unlike ungrammatical filler sentences, can satiate over repeated exposures; such a finding would not be expected if adjunct islands were in fact strong islands (i.e. a purely syntactic phenomenon).

As we further explained in section 3.5.1.2.3, Tucker et al.'s (2019) experimental investigation of RPs in MSA revealed that the strongest facilitation effect of RPs is found when RPs occur inside adjunct islands; however, no facilitation effect was found in complex NP islands (or relative clause islands). Tucker et al. attributed the lack of amelioration effect in CNPC islands to 'the unacceptability of any filler-gap dependency with CNPC constructions more generally' (2019, p.37). However, we assume that the amelioration effect of intrusive RPs in adjunct islands might occur due to the weaker nature of the violation of adjunct islands compared to RC islands.

4.2.3 Resumption and embedding

As illustrated earlier, resumptive pronouns tend to be more acceptable in deeply embedded structures (Erteschik-Shir, 1992; Asudeh, 2012). For instance, Tsimpli (1999) found that resumptive pronouns in Greek are acceptable when at least one CP-clause intervenes between it and its related filler phrase (5).

5.	Pion	ipoptefθike	i	Maria	oti	θa
	who.ACC	suspected.3SG	the.NOM	Maria	that	will
	ton	kalesume?				
	him.ACC	invite.1PL				

'Who did Maria suspect we will invite?'

(Greek; Tsimpli, 1999, cited in Alexopoulou and Keller, 2007, p.114)

Similarly, Sterian (2016) contended that RPs in bare *wh*-questions in Iraqi Arabic are allowed when they are deeply embedded, as in (6) and (7) below. Moreover, several corpus-based studies in true resumption languages have found that the use of RP

strategy increases as the integration site is more deeply embedded (Ariel, 1999; McCloskey, 2017).

6. a. Direct object extraction of bare interrogative – local

Iman minnu: ša:fat_____ bi-beyt Awatif

Iman who saw.3SF in-house Awatif

'Whom did Iman see at Awatif's house ?'

b. Direct object extraction with resumption

*Iman minnu: ša:fat=hu bi-beyt Awatif

Iman who saw.3SF=3MS in-house Awatif

*'Whom did Iman see [him] at Awatif's house ?'

7. a. Direct object extraction of bare interrogative – long distance

Suha minnu: taʕatagid ra:h yaʕzim_____ Ahmad ?

Suha who think.3FS will invite.3MS_____ Ahmad

'Whom does Suha think that Ahmad will invite_____ ?'

b. Direct object extraction with resumption

Suha minnu: taʕatagid ra:h yaʕzim=hu Ahmad ?

Suha who think.3FS will invite.3MS=3MS Ahmad

'Whom does Suha think that Ahmad will invite [him] ?'

We reasoned that dependencies with intrusive and obligatory RPs interact differently with embedding. If intrusive RPs are used to facilitate the processing of a demanding structure (Ariel, 1999; Hofmeister and Norcliffe, 2013; Chacón, 2019), we can expect their acceptability to be affected by the level of embedding (or distance). On the other hand, if obligatory true RPs are syntactically licensed, their acceptability should not be affected by the level of embedding. From a language processing point of view, however, long-distance dependencies are predicted to require a higher processing cost; this may result in a reduction in acceptability, even when true RPs are involved. In other words, resolving binding dependencies requires holding the filler phrase active in WM until the integration site is reached. According to WM-based accounts of FGD processing, holding filler phrases active in WM while processing clause

boundaries is expected to cause an overload in WM. Hence, we predict that grammatical RP dependencies in cases of cleft *wh*-questions and relative clauses are less acceptable in longer dependencies when compared to their shorter counterparts.

4.2.4 Resumption and WM

The literature on intrusive resumption has argued that the use of RPs is restricted to cases in which parsers are unable to hold filler phrases active in WM due to increased demands on WM resources (Ariel, 1999; Hofmeister and Norcliffe, 2013; Hammerly, 2019; Chacón, 2019). If this hypothesis is correct, we propose that working memory capacity will correlate with the ability to hold filler phrases active in long-distance dependencies and across island structures. Reliance on intrusive RPs in such structures might therefore be greater for individuals with lower WM capacity. Differences in WM capacity are not expected to affect the processing of true RPs, as they are grammatically licensed (i.e. they have no advantage in processing).

4.3 Hypotheses and predictions

The hypotheses that are drawn from the literature review are outlined below.

- Sensitivity to islands is a diagnostic for syntactic movement.
- Features of C affect the type of dependency; this implies that resumption behaviour differs across different types of dependency structures.
- True RPs appear in binding dependencies and fully restore their acceptability.
- Intrusive RPs appear in movement dependency and partially improve acceptability.
- The extent to which intrusive resumption can ameliorate islands varies across types of island: the literature indicates that the facilitation effect of RPs is more evident in weaker islands, like adjunct islands, than stronger islands like RC islands (Tucker et al., 2019).
- Individuals with lower WM capacities will be more accepting of intrusive RPs than individuals with higher WM capacities in cases of demanding dependencies (islands and long-distance dependencies).

Taking these hypotheses into account, we make the following predictions.

- 1) Only movement dependencies are sensitive to islands (i.e. bare *wh*-questions, but not relative clauses or cleft *wh*-questions).
- 2) RPs are obligatory in relative clauses and cleft *wh*-questions.
- 3) RPs are marginally acceptable in bare *wh*-questions:
 - a. to alleviate island violations;
 - b. to facilitate the processing of long-distance dependencies.
- 4) The amelioration effect induced by RPs is greater in adjunct islands than in RC-islands.
- 5) The acceptability of RPs (vs gaps) is affected by the distance of the dependency only if syntactic movement is involved (i.e. in *wh*-questions, but not in relative clauses or *wh*-clefts).
- 6) Long-distance dependencies yield a marginal decrease in acceptability:
 - a. of binding structures with true RPs; and
 - b. of movement structures with gap.

4.4 Experiment

4.4.1 Methodology

4.4.1.1 Design of the experiment

The experiment has been designed to examine the interaction between the following factors:

- Dependency structures:
 - o Relative clause
 - o Clefted *wh*-questions
 - o Bare *wh*-questions
- Tail:
 - o RP
 - o Gap
- Condition:

- *Baseline*: integration site is in the matrix clause
- *Crossed-non-island*: integration site is in an embedded non-island clause
- *Crossed adjunct-island*: integration site is in an embedded adjunct island clause
- *Crossed RC island*: integration site is in an embedded RC island clause

Crossing these factors yielded the following conditions in each type of *WH*-dependencies:

- Baseline – Gap/RP
- Crossed non-island – Gap/RP
- Crossed adjunct island – Gap/RP
- Crossed RC-island – Gap/RP

Several issues have been taken into account when designing the experiment. First, only dependencies involving direct objects were included in the experiment. This is due to the fact that overt resumptive pronouns do not appear at the highest subject positions, either due to the Highest Subject restriction (Shlonsky, 1992; McCloskey, 2002) or due to the pro-drop property of Arabic (Soltan, 2007); in Arabic, there must be an agreement marker on the verb in cases of subject fillers in relative clauses and clefted *wh*-questions, which might indicate the presence of a null resumptive pronoun (Soltan, 2007). However, no such agreement marker appears on the verb when the grammatical role of the filler in the dependency structure is that of an object, which ensures the existence of a genuine gap rather than a null resumptive pronoun in the extraction site.

Second, the present experimental study controlled the role of animacy and the discourse-linked (*D-linked*) properties of filler phrases by investigating dependencies formed only with inanimate non-D-linked fillers. This is because the literature on resumption indicates that resumption is sensitive to the animacy and d-linking properties of filler phrases (Tsimplici, 2003; Aoun et al., 2009; Alexopoulou and Keller, 2013); see (8) and (9). Hence, we suggest that investigating resumption with inanimate fillers will facilitate a better investigation of the interaction between resumption and locality factors, ensuring that any effect of resumption is independent from the semantic or pragmatic features of filler phrases.

8. a. *šu štarayt-i
 what bought.2sf-it
 ‘What did you buy?’
- b. miin/?ayya maariḍ zarit-u naadia?
 who/which patient visited-3SG.F-3SG.M.OBJ Nadia
 ‘Who/which patient did Nadia visit?’
 (Syrian Arabic; Sulaiman, 2016)
9. a. Pjon ipes oti (ton) prosevalan xoris
 Whom said-2S that him-insulted-3p without
 logho?
 reason?
 ‘Who did you say that they insulted (*him) without reason?’
- b. *Ti nomizis oti tha to dhiavasun?
 What think-2S that will it read-2
 ‘What do you think that they will read?’
 (Tsimpli, 2003, p.217)

Third, all experimental items are constructed in such a way that the filler phrases always head bi-clausal dependencies, ensuring that acceptability will not be affected by the length of the dependency structure. For illustrative purposes, the schematics of experimental sentences are presented in (10) for bare and cleft-*wh*-questions and (11) for relative clauses.

10. Bare *wh*-questions and cleft-*wh*-questions:

[_{CP}Filler [_{CP}?]].

11. Relative clauses:

[_{CP} [_{CP}Filler [_{CP}]].

We notice that relative clauses differ in that the dependency structure is embedded inside a main clause. This was necessary so that the relative clauses would sound plausible. However, similar to the experimental items of bare *wh*-questions and cleft

wh-questions, the dependency structure itself is bi-clausal in relative clauses. Thus, across experimental sentences, all dependency structures are bi-clausal, but differ in terms of whether or not a dependency crosses a clause boundary.

The requirement that all experimental items be bi-clausal initially led us to add a second clause to the baseline condition, following the matrix clause. However, we found that creating such sentences with an embedded non-island clause is not possible in Baha Arabic when the filler phrase is object and inanimate. Accordingly, to create bi-clausal experimental sentences for this baseline condition, we included a non-crossed embedded ‘adjunct’ island structure, as illustrated in the example below.

12. **ʔayš** amal zar-at-/=h lamman kan-at fi
What amal visit-3SF-/=it when was-3SF in
ar-ryaḍ?
the-Riyadh?
‘What did Amal visit (it) [when she was in Riyadh?]

Hence, the baseline condition is referred to as the ‘non-crossed island condition’. Thus, the manipulated conditions across the different dependency structures are updated as follows:

- Non-crossed island – Gap/RP
- Crossed non-island – Gap/RP
- Crossed adjunct island – Gap/RP
- Crossed RC-island – Gap/RP

Finally, sentences across the different conditions are not lexically matched (except for RP conditions vs gap conditions). In other words, sentences across conditions are presented in minimally different pairs in which the type of element appearing at the dependency tail (i.e. gap vs RP) is manipulated. This is done purposely to avoid any potential effect of lexical priming effect (as we are utilising a repeated measures design). Using lexically matched sentences across the different variables (Tail, Structure, and Condition) would have required each lexicalisation to be repeated 24 times, which would have induced an accommodation effect. The use of different lexicalisation sets across conditions will be taken into consideration during statistical analysis.

4.4.1.2 Materials

We systematically crossed the factors of Condition (of the embedded clause), Tail (type) and dependency Structure (type) in a 4 x 2 x 3 design, resulting in 24 conditions (eight per dependency type); each of these was then lexicalised six times, resulting in 144 critical items. Using the Latin Square Method, these items were split into six randomised, counter-balanced lists, which were in turn used to pseudo-randomise the test items. Each list consists of 24 experimental items (only one sentence from each condition) and 48 fillers (2:1 filler to experimental ratio). The structure of the filler sentences will be described below. Distributing the experimental items into six lists ensured that only one sentence from each condition occurred in each list; see Table 4-1 for more details.

This study used a repeated measures design, meaning that each participant was asked to rate the acceptability of all sentences in all conditions (all participants were presented with all six lists). Due to the very large number of items participants were required to rate (144+288), experimental items were distributed across two surveys, each consisting of three blocks of 72 items. Each block represents one of the six lists created using the Latin Square Method. Hence, three sentences from each condition were presented in each survey. The method used to construct the blocks for each experiment ensured that the experimental items were not composed of identical lexical items. Thus, sentences of each pair of lexicalisations appear in different surveys. The interval between completing the two surveys was one week.

Table 4-1 Conditions involved in the first experiment.

Structure		Bare <i>wh</i> -question	Relative clause	Clefted <i>wh</i> -question
Condition of embedded clause	Tail			
Non-Crossed Island	Gap/RP			
Crossed Non-Island	Gap/RP			
Crossed Adjunct Island	Gap/RP			
Crossed RC-Island	Gap/RP			

Examples of experimental items for each of the different conditions in the three different dependency structures utilised in the experiment are presented below.

13. Bare *WH*-questions:

a. Non-Crossed Island, Gap/RP

ʔayš amal zar-at-/uh lamman kan-at fi
What Amal visited-3SF-/it when was-3SF in
ar-ryaḍ?
the-riyadh?
'What did you visit (it) when you were in Riyadh?'

b. Crossed Non-Island, Gap/RP

ʔayš tetwaqʕ-een ʔnn mħmmad raḥ yebeeʕ-/uh?
what think-2SF that Muhammad will sell.3SM-/it?
'What do you think that Muhammad will sell? '

c. Crossed Adjunct Island, Gap/RP

ʔayš muna ḥaḍar-at al-muḥaḍarah qablma taqrʔ-/uh?
what Muna attended-3SF the-lecture before read.3SF -/it?
'What did Mona attended the lecture before she read? '

d. Crossed RC-Island, Gap/RP

ʔayš mħmmad ʕazam al-fannan illi
what Muhammad invited.3SM the-artist who
rasam-/uh?
drew.3SM -/it?
'What did Muhammad invited the artist who drew? '

14. Clefted *WM*-questions:

a. Non-Crossed Island, Gap/RP

ʔayš al-fustan illi lebes-ti-/=h lamman ru-ḥti
What the-dress that wore-2SF-/=it when attended-2SF
al-ḥaflah?
the-party?

‘What is the dress that you wore when you attended the party?’

b. Crossed Non-Island, Gap/RP

ʔayš as-syarah illi qul-ti ʔnn ahamad
What the-car that said-2SF that Ahmad
sadam-/=ha?
hit.3SM-/=it?

‘What is the car that you said that Muhammad hit?’

c. Crossed Adjunct Island, Gap/RP

ʔayš al-maqal illi bent-k faz-at
What the-article that daughter-your won-3SF
bi-al-musabaqah lamman qarʔ-at-/=uh?
by-the-competition when read-3SF -/=it?

‘What is the article that your daughter won the competition when she read?’

d. Crossed RC-Island, Gap/RP

ʔayš al-jihaz illi al-mudeer karram
What the device that the-administrator honoured.3SM
al-moazaf illi šanaʕ -/=uh?
the-employer who made.3SM-/=it?

‘What is the device that the administrator honoured the employer who made?’

15. Relative clauses:

a. Non-Crossed Island, Gap/RP

ḥabai-t aš-šuqah illi ʔstʔjar-ti-/=ha lamman
liked-1S the-flat that rented-2SF -/=it when
ruḥ-ti Turkia
went-2SF Tutkey

'I liked the flat that you rented when you went to Turkey.'

b. Crossed Non-Island, Gap/RP

šef-t as-saḥah illi qul-ti l-ee ʔnn saleḥ
saw-1s the-watch that told-2SF for-me that Saleh
ʔštra--/ha
bought.3SM-/=it

'I saw the watch that you told me that Saleh bought.'

c. Crossed Adjunct Island, Gap/RP

kabbai-t al-ḥaleeb illi khaled meriḍ
spilled-1S the-milk that Khaled felt-sick.3SM
baḥd-ma šerib -/=uh
after drank.3SM-/=it

'I spilled the milk that Khaled became sick after he drank.'

d. Crossed RC-Island, Gap/RP

ʔkal-t al-ḥala illi muna teḥref al-bent illi
ate-1S the-dessert that Muna knew.3SF the-girl that
saww-at-/=uh
made-3SF-/=it

'I eat the dessert that Mona knows the girl who made .'

As for the filler sentences, these are meant to match with the experimental items as much as possible in terms of number of clauses and overall length. These included sentences that are both perfectly acceptable and completely unacceptable; this was done in order to encourage the participants to utilise the full range of the rating scale, and hence avoid scale bias. Half of the fillers were questions and half were declarative sentences. An example of a perfectly acceptable filler sentence is presented in (16). The unacceptable filler sentences included various grammatical violations, such as subject-verb agreement mistakes (17), sub-categorisation errors (18), and *wh*-questions with incorrect *wh*-words (19).

16. Mona qalat lee ?enn-ha begyet fii al-mustashfa lhad
Mona told me that-she stayed in the-hospital till
al-fajr
the-morning
'Mona told me that she stayed in the hospital till the morning.'

17. Laish almdarresat maša l-beaut-hum
Why the teacher.PF went-3SM to-house-their
badree?
early?

18. Layla qalat inn Mohammad yeŕtaqed al-kitaab
Layla said that Mohammad thinks the-book.

19. Meen akalt-ii lamman reḥti l-al-maṭŕam
Who ate.2F when went.2F to-the-restaurant
al-jedeed?
the-new?

Regarding the working memory task, the backward digit span task was utilised for this purpose. In this task, participants were asked to repeat backwards sequences of numbers with increasing length that had been presented to them on a computer screen. For this study, the backward digit span task was implemented as described by Woods et al. (2011); this is a newer approach that, unlike the traditional method, does not require each length to be presented twice, nor for the test to end after two

errors being made at the same length. In this task, each participant was required to complete 14 trials, beginning with length 2. If the participant remembered a length correctly, they could move on to the next length; if they made an error, they were given another trial of the same length. If they made two errors in the same length, the test did not end, but the participant was moved down to a lower length. The test ended after 14 trials.

4.4.1.3 Subjects

For this study, 39 native speakers of Baha Arabic were recruited (through email invitations) to complete the acceptability judgement task. Nineteen of the subjects were male, while twenty were female; all participants were within the 18-40 age bracket. Participants completed a language background questionnaire to ensure that they had similar linguistic profiles (in terms of having the same level of bilingualism and speaking the same variety of Arabic, specifically Baha Arabic), as this would ensure that their responses would not be influenced by their knowledge of other languages. The language background questionnaire indicated that all participants were native speakers of Baha Arabic. (A translated version of this questionnaire is attached in the Appendix.)

4.4.1.4 Procedures

In order to satisfy the ethical standards of research, full ethical approval was obtained prior to commencing the experiment (reference number: LTSLCS-039). Participants were asked to read an information sheet and provide their informed consent before taking part in the study.

After completing a brief, online backward digit span task, participants were asked to complete the first online questionnaire. In this questionnaire, participants were asked to rate the acceptability of the sentences presented to them on the screen, one at a time. The second questionnaire was sent to the participants one week after they had completed the first one. Participants completed the two questionnaires in locations of their own choice, as they were both conducted online. Neither questionnaire took longer than 25 minutes in total to complete.

Before beginning the acceptability judgement task, subjects read a written description of the task and instructions on how to complete it. They were then asked to indicate on a scale from 1 to 7 how acceptable they thought each sentence was, with 7 being

‘completely unacceptable’, and progressively lower numbers representing increasingly higher acceptability (with 1 being ‘completely acceptable’).

In order to encourage the participants to use the entire range of the scale, which would minimise scale bias (Sprouse and Schütze, 2017), the fillers also included anchor points for ‘perfectly good’ and ‘totally unacceptable’ sentences. Participants were asked to take a break of about five minutes between each survey block.

Of the total participant group, nine participants completed only the first questionnaire; as a result, they only rated three items per condition. The remainder (N=30) rated all items (i.e. six per condition).

4.4.2 Results

4.4.2.1 Calculating WM scores

As noted above, the backward digit span task used in the present research was the same as that described by Woods et al. (2011). Woods et al.'s (2011) experimental study, which aimed to test the reliability of backward digit span tasks, found that the traditional method of scoring these tasks was not reliable. The traditional method depends on two scores, specifically two-error maximum length (TE-ML) and two-error total trials (TE-TT); however, these authors' results revealed that the TE-ML metric underestimated the true ML span of subjects by more than 0.5 digits, while the TE-TT metric was more unreliable still, as it ‘showed higher variance, a greater coefficient of variation, poorer test-retest reliability, [...] and poorer correlations with scores on other neuropsychological tests of memory’ (Woods et al., 2011, p.12).

Woods et al. (2011) accordingly argue in support of two other metrics: the ML metric, defined as the maximum length correctly reported over all 14 trials, and the MS metric, referring to the mean span over 14 trials. When comparing the reliability of the scores of these two metrics, Woods et al. (2011) eventually argued in support of the MS metric due to ‘its insensitivity to the number of lists presented’ (Woods et al., 2011, p.12). In light of these findings, the present study will score the participants' WM capacities using the MS metric. To illustrate the scoring procedures, the results of a backward digit span test for one subject are presented in Table 4-2 below.

Table 4-2 Results of the backward digit span task of one subject.

The first column of the table illustrates the number of trials presented to the participant. The result column denotes the accuracy of the participant's responses (0=wrong; 1=correct).

Trial	Length	Presented digits	Response	Result
1	2	93	93	1
2	3	129	129	1
3	4	.4867	4867	.1
4	5	94879	9487	.0
5	5	61768	76	0
6	4	1547	1547	1
7	5	.61784	61784	.1
8	6	.625426	625429	0
9	6	169253	169253	.1
10	7	4618396	4618396	1
11	8	91381915	91381915	1
12	9	352192849	352192849	1
13	10	1832927525	832927525	0
14	10	9379712952	54367	0

The list length of each trial, presented in the second column, increased by one before the subject's first mistake at length 5 (trial 4). Following that, list lengths were varied between 4 and 10 digits. The table shows that the subject continued to achieve an ML of 9 (trial 12). The MS was 8.33 as calculated by adding the hit rate for each list length (e.g. 2 = 1.0, 3 = 1.0, 4 = 1.0, 5 = 0.33, 6 = 0.5, 7 = 1.0, 8 = 1.0, 9 = 1.0, and 10 = 0.0, sum = 6.08) to the baseline value of 1.5.

4.4.2.2 Descriptive data

We begin with an overview of the descriptive results of the acceptability data. As outlined above, the factors we manipulated in the present study are as follows: Condition [of embedded clause] (NonCrossed-Island/Crossed-NonIsland/Crossed Adjunct-Island/Crossed RC-island), Tail (RP, Gap) and Structure (Bare-WH/RC/Cleft-WH). We are primarily interested in investigating the interaction of RPs with the different conditions across the three dependency structures in the present study. We hypothesised that, if the different structures have different underlying syntactic structures (binding vs movement), they will behave differently in terms of their interaction with the condition and resumption. Figure 4-1 illustrates the descriptive statistics related to participants' judgements for the experimental sentences.

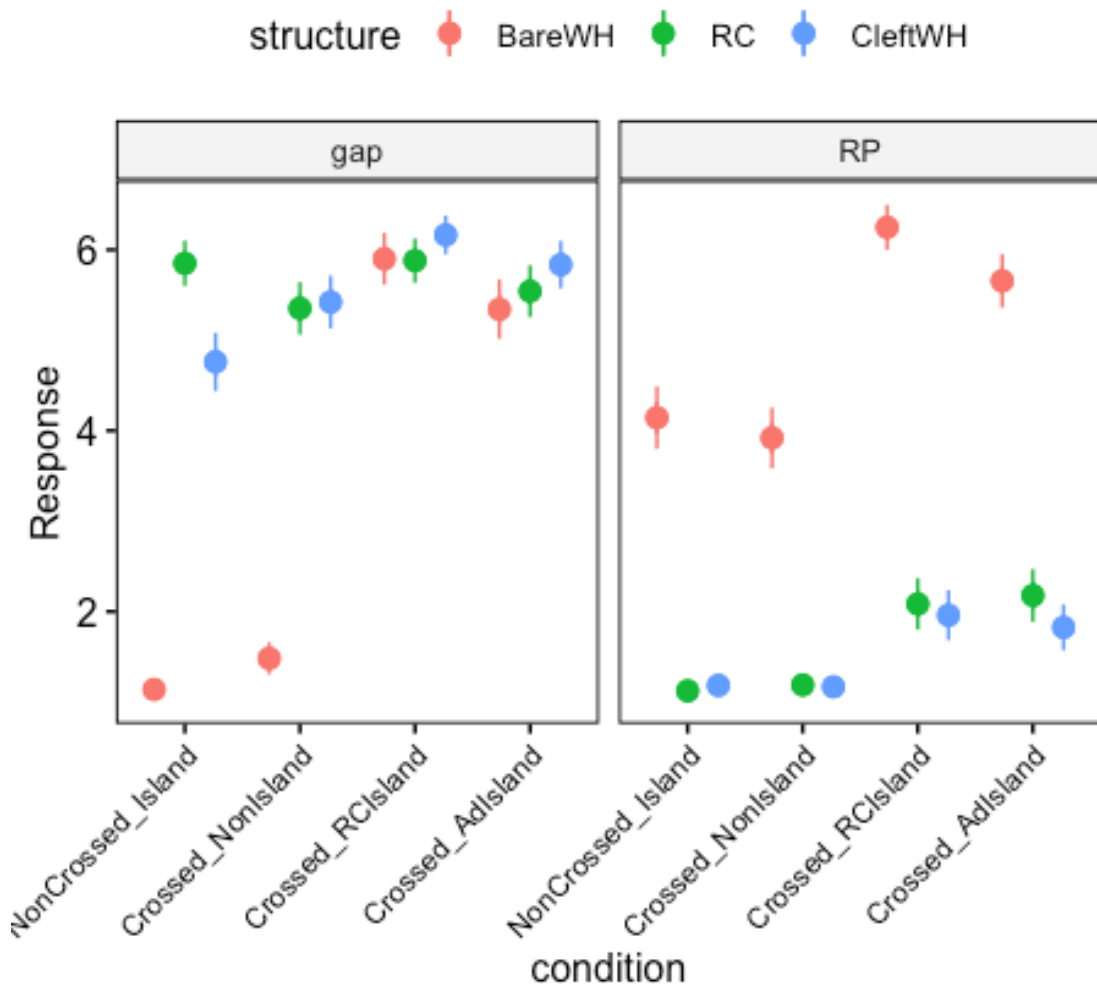


Figure 4-1 Rejection ratings showing mean and 95% confidence intervals across conditions and dependency structures, with and without RP.

Rejection scores range from 1 ('completely acceptable') to 7 ('completely unacceptable').

From visual inspection of the data, it appears that Clefted-WH and RC sentences, which both feature the relative complementiser *illi*, appear to exhibit a similar pattern, except in the NonCrossed-Island condition, where Clefted-WH with gaps are rejected less than RC with gaps, although both types of structures are in the unacceptable range (Mean, Clefted-WH: 4.69; RC:5.74).

Moreover, there is no marked difference between Island types within each dependency structure, albeit with one possible exception in the case of bare *wh*-questions featuring RPs (in this case, Adjunct islands appear to be rejected marginally less than Relative Clause islands).

We accordingly decided to treat islands as a single category in the analyses, and further to limit the comparison of structures to two (structures featuring *illi* (i.e. relative clauses and clefted *wh*-questions) vs structures not featuring *illi* (i.e. bare *wh*-questions)). Consequently, the number of conditions was reduced from 24 to 12, resulting from the crossing of three variables with 3 x 2 x 2 levels:

- Condition (Crossed Island/Crossed Non-Island/Non-Crossed Island);
- Tail (RP/gap);
- Structure (*wh*/*illi*-structures).

Figure 4-2 illustrates the descriptive statistics related to participants' judgements of experimental items after collapsing the variables. Higher numbers on the seven-point scale represent worse judgements; 7 represents 'completely unacceptable', while progressively lower numbers represent increasingly favourable judgements (with 1 being the highest rating of 'completely acceptable').

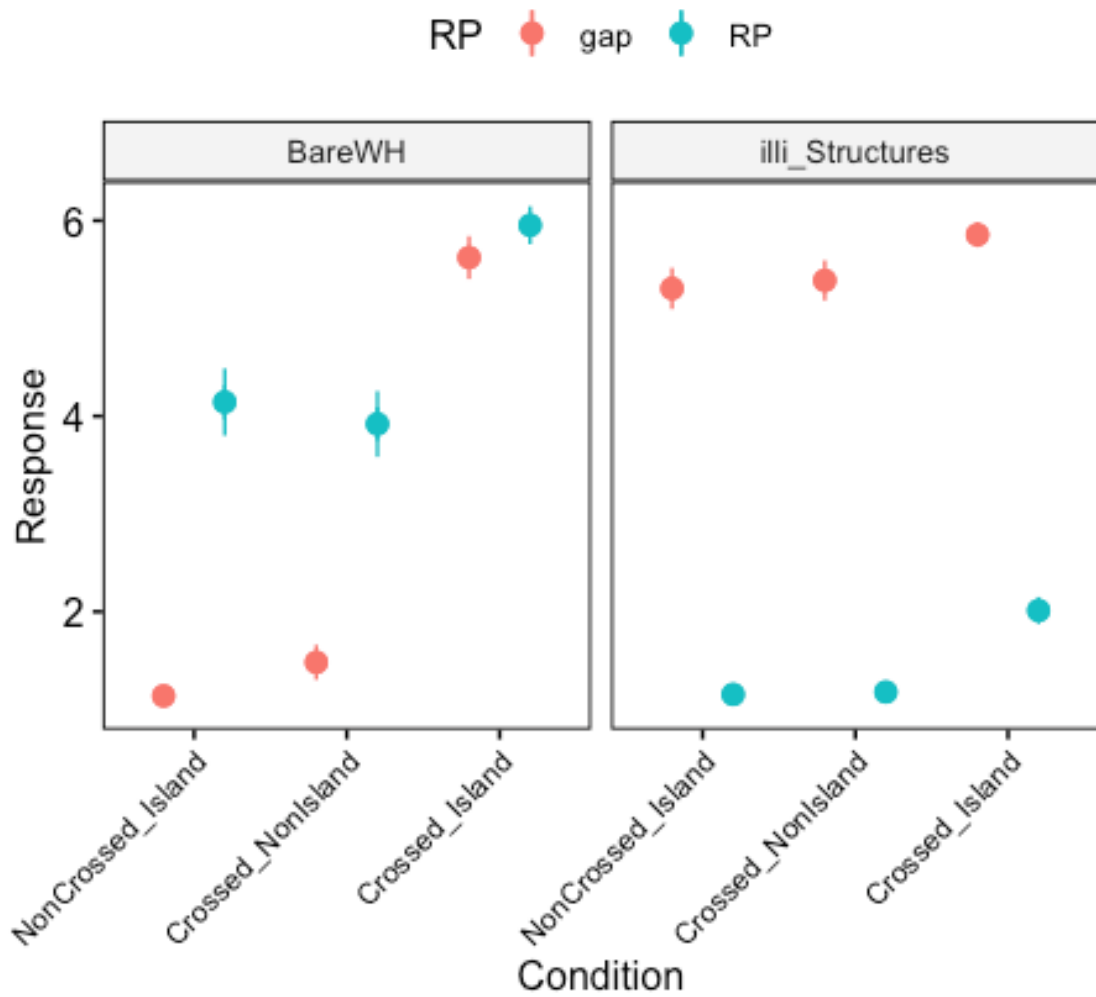


Figure 4-2: Mean rejection ratings with confidence intervals for the different conditions, each with and without RP, across the three dependency structures.

Higher numbers on the 7-point scale represent worse judgements; 1 = 'completely acceptable', 7 = 'completely unacceptable'.

This visualisation of raw data suggests the existence of a three-way interaction between Structure, Condition and Tail. More specifically, it appears that bare *wh*-questions and *illi*-structures behave differently in terms of the way they interact with resumption across the different conditions.

Table 4-3 presents a comparison between the mean rejection of the different conditions included in this experiment and the mean rejection of (grammatical vs ungrammatical) filler sentences.

Table 4-3 Comparison between mean rejection of core experimental items with fillers.

Structure	Condition	RP	Response.mean	Response.sd
RC_CleftWH	Crossed_NonIsland	RP	1.18	0.81
RC_CleftWH	Crossed_NonIsland	Gap	5.39	2.09
RC_CleftWH	Crossed_Island	RP	2.01	1.99
RC_CleftWH	Crossed_Island	Gap	5.85	1.83
RC_CleftWH	NonCrossed_Island	RP	1.15	0.71
RC_CleftWH	NonCrossed_Island	Gap	5.30	2.14
BareWH	Crossed_NonIsland	RP	3.92	2.42
BareWH	Crossed_NonIsland	Gap	1.48	1.28
BareWH	Crossed_Island	RP	5.95	1.98
BareWH	Crossed_Island	Gap	5.62	2.23
BareWH	NonCrossed_Island	RP	4.14	2.48
BareWH	NonCrossed_Island	Gap	1.14	0.66
Filler sentences	Response.mean		Response.sd	
Grammatical	1.28		1.05	
Ungrammatical	5.96		1.87	

4.4.2.3 Methods of analysis

In principle, it is possible that the different lexicalisation sets across conditions might have introduced a confounding factor into the analysis. We thus carry out two analyses to mitigate against this. First, we use acceptability rating as the dependent variable and include lexicalisation as a random effect in the model (where each level

corresponds to a pair of sentences: with gap or with RP). Inspecting the distribution of the random effect values for lexicalisation across the different conditions and structures will bring any confounding effects to light. In the second analysis, we use the differential score [Gap (rating) – RP (rating)] as the dependent variable and include lexicalisation as a random effect (with each level corresponding to a single sentence). This allows us to investigate how the size of the difference between the acceptability of gap dependencies and RP dependencies in each minimal pair for each participant is affected by the condition of the embedded structure across the different types of *WH*-dependencies.

We will begin below with the analysis in which rating score is used as the dependent variable.

4.4.2.4 Statistical analysis

4.4.2.4.1 Rating score as the dependent variable

The judgement data in this study was collected using an ordinal scale, where participants' scores of acceptability judgements were recorded on a seven-point scale. As this is an ordinal scale, it was likely that some participants would interpret the individual levels differently than others (for example, some subjects might tend to make only extreme ratings on the scale, while others will tend to avoid using these extreme points while rating). Hence, to force participants to make use of the rating scale's full range, the filler sentences added to the survey included both fully acceptable and fully ungrammatical items, as described above.

However, there are still probable variations among participants, which can only be partially controlled by statistical models. In particular, the distance between the levels of response scale might not be perceived as identical across participants. In order to take this into account, we fitted the analysis using linear mixed effects ordinal regression models with a cumulative link function (to allow the use of random effects). This was done in R (version 3.5.0) using the package *ordinal* (version 2019.3-9) (Christensen, 2015; R Core Team 2017). This type of model assumes that the individual points on the scale are ordered, without assuming equidistance between the points on the scale.

Modelling procedure

The models in this analysis are fitted bottom-up, starting from random effects only (as a null hypothesis model) and incrementally adding fixed effects. The first model was constructed with Subject and Lexicalisation as random factors. The basic model was subsequently extended to include one variable each time, using likelihood ratio comparisons to ascertain whether adding the variable significantly improved model fit (and discarding it otherwise).

Prior to the statistical analysis, one participant was excluded for not having performed the WM task. The WM variable was scaled and centred to facilitate model interpretation.

Likelihood ratio comparisons between nested models revealed that the optimal model was the one including main effects and three-way interaction of fixed effects (Structure, Tail, Condition) and an interaction between (centered.WM, Tail, Condition and Structure), as well as a random intercepts for Lexicalization and Subject, and random slopes for Subject by Tail, Subject by Condition, Subject by Tail in interaction with Condition, and random slope for Lexicalization by Tail. A summary of the optimal model is presented in Table 4-4. The summary of random effects coefficients is presented in Table 4-5, while threshold coefficients are summarised in Table 4-6.

Table 4-4 Summary of the optimal Ordinal Cumulative Link Mixed Model (clmm) model of the Rejection ratings.

The formula of the model is as follows: `clmm(as.factor(response) ~ (Condition * Tail * Structure) + (centered.WM:Condition:Tail:Structure) + (1|subject) + (1|Lexicalisation) + (1|Lexicalisation:Tail) + (1|subject:Condition) + (1|subject:Tail) + (1|subject:RP:Condition), data=datq)`. Reference levels: WH/gap/NonCrossed_Island.

Coefficients	Estimate	Std.Error	z-value	p-value
Condition				
NonCrossed_Island	—			
Crossed_NonIsland	1.1	0.52	2.17	0.030
Crossed_Island	6.5	0.48	13.41	<0.001
Condition *Tail				

Crossed_Island * RP	-4.2	0.48	-8.82	<0.001
Crossed_NonIsland * RP	-1.5	0.52	-2.79	0.005
Condition * Tail * Structure				
Crossed_Island * RP * illi-structures	5.7	0.58	9.95	<0.001
Crossed_NonIsland * RP *	1.4	0.66	2.08	0.038
Condition * Structure				
Crossed_Island * illi-structures	-5.8	0.53	-10.95	<0.001
Crossed_NonIsland * illi-structures	-0.95	0.59	-1.62	0.11
Tail				
Gap	—			
RP	4.6	0.44	10.4	<0.001
Tail * Structure				
RP * illi-structures	-10	0.52	-19.66	<0.001
Structure				
BareWH	—			
illi-structures	5.6	0.46	12.13	<0.001
Condition * Tail * Structure * centered.WM				
Crossed_Island * gap * BareWH * centered.WM	0.08	0.10	0.82	0.4
Crossed_Island * gap * illi-structures * centered.WM	-0.27	0.09	-3.07	0.002

Crossed_Island * RP * BareWH * centered.WM	0.00	0.10	0.01	>0.9
Crossed_Island * RP * illi-structures* centered.WM	-0.06	0.09	-0.70	0.5
Crossed_NonIsland * gap * BareWH * centered.WM	-0.05	0.13	-0.39	0.7
Crossed_NonIsland * gap * illi-structures* centered.WM	-0.19	0.09	-2.04	0.042
Crossed_NonIsland * RP * BareWH * centered.WM	-0.32	0.10	-3.17	0.002
Crossed_NonIsland * RP * illi-structures* centered.WM	-0.08	0.14	-0.60	0.5
NonCrossed_Island * gap * BareWH * centered.WM	-0.13	0.18	-0.69	0.5
NonCrossed_Island * gap * illi-structures * centered.WM	-0.17	0.09	-1.86	0.062
NonCrossed_Island * RP * BareWH * centered.WM	-0.11	0.10	-1.11	0.3
NonCrossed_Island * RP * illi-structures * centered.WM	-0.28	0.16	-1.72	0.085

Table 4-5 Coefficients of random effects

Groups	Variance	Std.Dev.
subject:RP:Condition	0.14	0.4
Lexicalisation:RP	0.14	0.4

subject:Condition	0.05	0.23
subject:RP	0.55	0.74
Lexicalisation	0.3	0.54
Subject	0.3	0.53

Table 4-6 Threshold coefficients.

	Estimate	Std.Error	p value
1 2	3.4	0.44	<0.001
2 3	3.9	0.44	<0.001
3 4	4.2	0.44	<0.001
4 5	4.6	0.44	<0.001
5 6	5.0	0.44	<0.001
6 7	5.7	0.45	<0.001

For ease of interpretation, the coefficients of the model are plotted in Figure 4-3. This figure visualises the probability of each rating across conditions, rather than the mean rating. Each vertical line corresponds to a combination of factor levels arising from the interaction between Structure, RP and Condition (as determined by the optimal model). The points at which this vertical line crosses the coloured lines indicates the probability of each rating (on the y-axis) for that combination of factor levels. The levels of Condition are abbreviated as follows: Non-Crossed Island (NoCI) / Crossed Non-Island (CNoI) / Crossed-Island (CI).

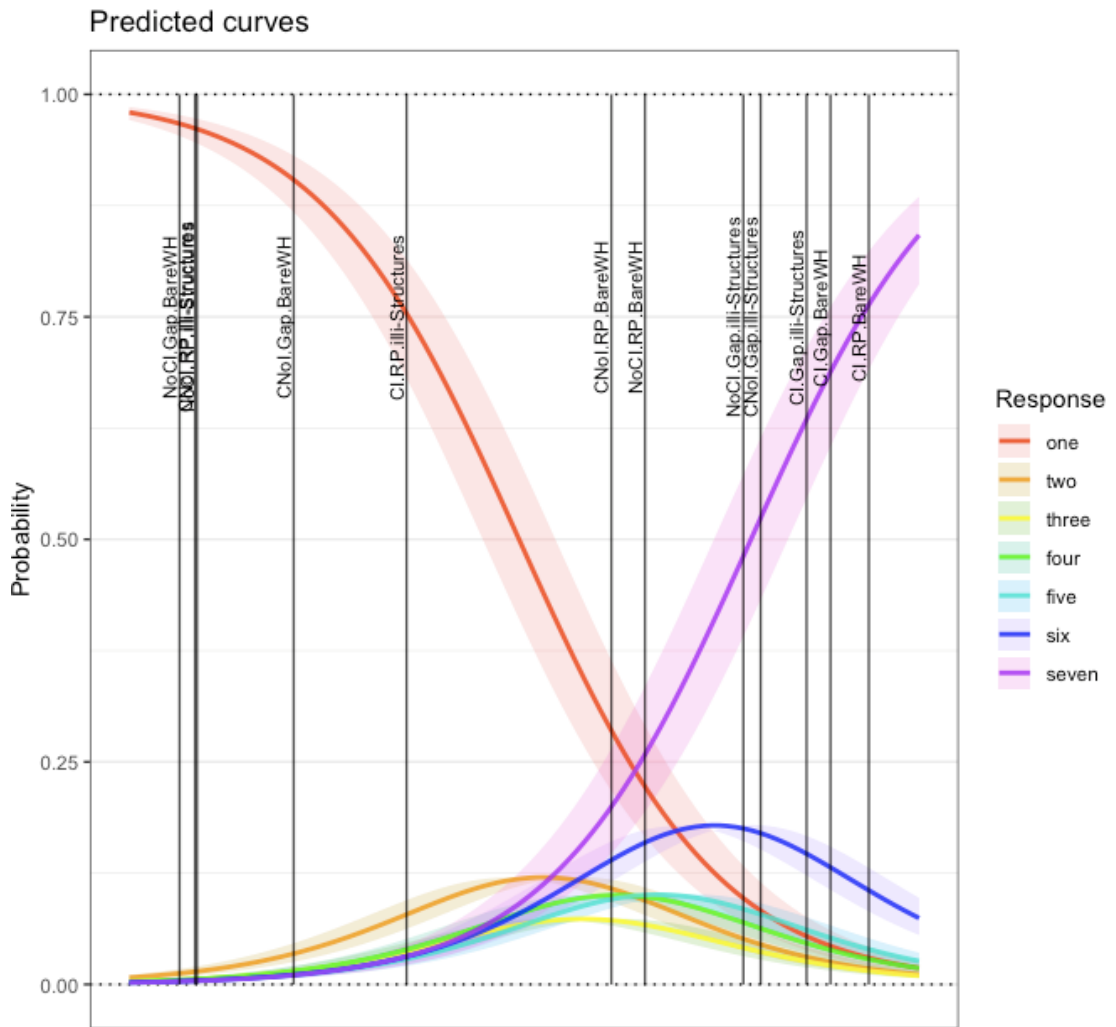


Figure 4-3 Results of the Ordinal Cumulative Link model showing the probability curves for each of the seven ratings in each condition (represented with vertical lines), along with Confidence Intervals.

The purple line (7) represents the lowest level of acceptability rating, while the red line (1) represents the highest level of acceptability.

As can be seen from Figure 4-3, the pattern of acceptance rate differs across conditions, as follows:

- Some conditions consistently yielded very high acceptance ratings. In four conditions, the probability of full acceptance was higher than 85%: specifically, Non-Crossed Island and Crossed Non-Island with gaps in bare *wh*-questions, and Non-Crossed Island and Crossed Non-Island with RPs in *illi*-structures. In *illi*-structures that crossed an island boundary and featured a RP, the likelihood of full acceptance dropped to 75%.

- Two conditions yielded mixed ratings, with equal probability (25% each) of being either highly accepted or highly rejected: specifically, Non-Crossed Island and Crossed Non-Island with RPs in bare *wh*-questions.
- Five conditions yielded high levels of rejection, but with various levels of probability (ranging from about 60% chance of a 7 or 6 score in gapped relative clauses, to a 75% chance of a 7 score in bare *wh*-questions crossing an island boundary).

The pattern of acceptance rate across the experimental conditions revealed a significant three-way interaction between the factors Tail, Structure and Condition. Furthermore, this pattern of interaction differs across the two types of dependency structures.

Starting with bare *wh*-questions, acceptability patterns are mainly determined by the Condition of the embedded structure, with the Crossed Island condition receiving significantly higher rates of rejection than the Non-Crossed Island condition (Estimate: 6.5, z : 13.41, p : <0.001) and Crossed Non-Island condition (Estimate: 5.35, z : 14.07, p : <0.0001). This higher level of rejection reflects the effects of violating island constraints. Furthermore, the results revealed a significant difference between the acceptability of Non-Crossed Island and Crossed Non-Island conditions (Estimate: 1.13, z = 2.17 p = 0.03), suggesting the existence of an embedding effect. Moreover, the results revealed a main effect of Tail, with gaps being preferred to RPs (Estimate: 4.6, z : 10.4, p : <0.001).

Experimental results also revealed an interaction between Tail and Condition: although gaps are preferred to RPs across the different conditions, the size of this effect across conditions is not consistent. In the Crossed Island condition, RPs are as bad as gaps; a pair-wise comparison test revealed no significant difference between gaps and RPs in this condition (Estimate: 0.4, z : 1.3, p : 0.2). In other words, RPs do not significantly ameliorate island effects.

In the Non-Crossed Island condition and Crossed Non-Island conditions, RPs received an inconsistent rating pattern. The results clearly revealed that RPs are rejected more than gaps in the Non-Crossed Island condition (Estimate: 4.6, z : 10.4, p : <.001), and in the Crossed Non-Island condition (Estimate: 3.17, z : 8.39, p : <.0001). However, a pairwise comparison test revealed no significant difference between the acceptability of Non-Crossed Island with RP and Crossed Non-Island with RP (Estimate: 0.322, z : 0.41, p : 0.785), suggesting that RPs are not any more accepted

than gaps when the *wh*-dependency structure crosses a non-island clause boundary. In summary, RPs are never rated higher than gaps in *wh*-questions, regardless of the length of the dependency and the crossing of an island boundary. If no island boundary is crossed, gaps are clearly preferred.

Furthermore, careful investigation of participants' ratings for Non-Crossed Island and Crossed Non-Island conditions with RPs revealed that the gradient nature of their pattern of acceptability does not reflect a consistent intermediate rating. We found that the gradient nature of this pattern of acceptability reflects significant inter-individual variation. Figure 4-4 suggests the existence of strong variability within participants; in short, the same participants sometimes rated sentences in this condition as acceptable, sometimes unacceptable, and sometimes tended to choose the middle point of the scale. It would further seem that this variability is not due to a lexicalisation effect, as it can be observed from Figure 4-5 that experimental sentences of different lexicalisation sets received an inconsistent rating pattern. Figure 4-6 further shows that variability within participants is not conditioned by lexicalisation.

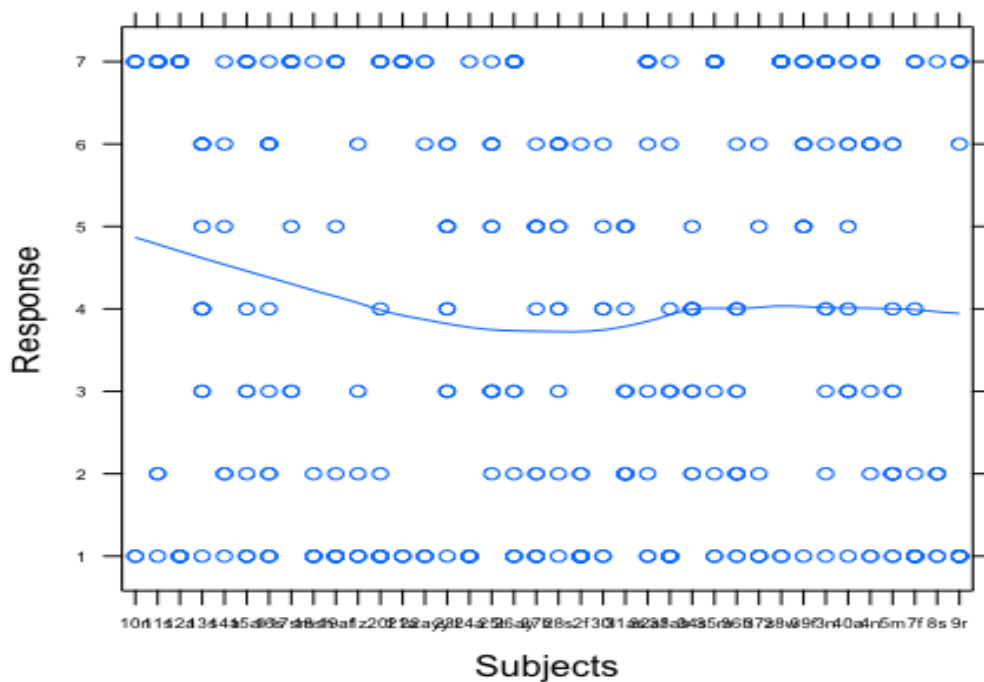


Figure 4-4 Variability within participants' rating of bare *wh*-questions with RPs in the NonCrossed_Island and Crossed-NonIsland conditions.

Ratings are not affected by Subjects.

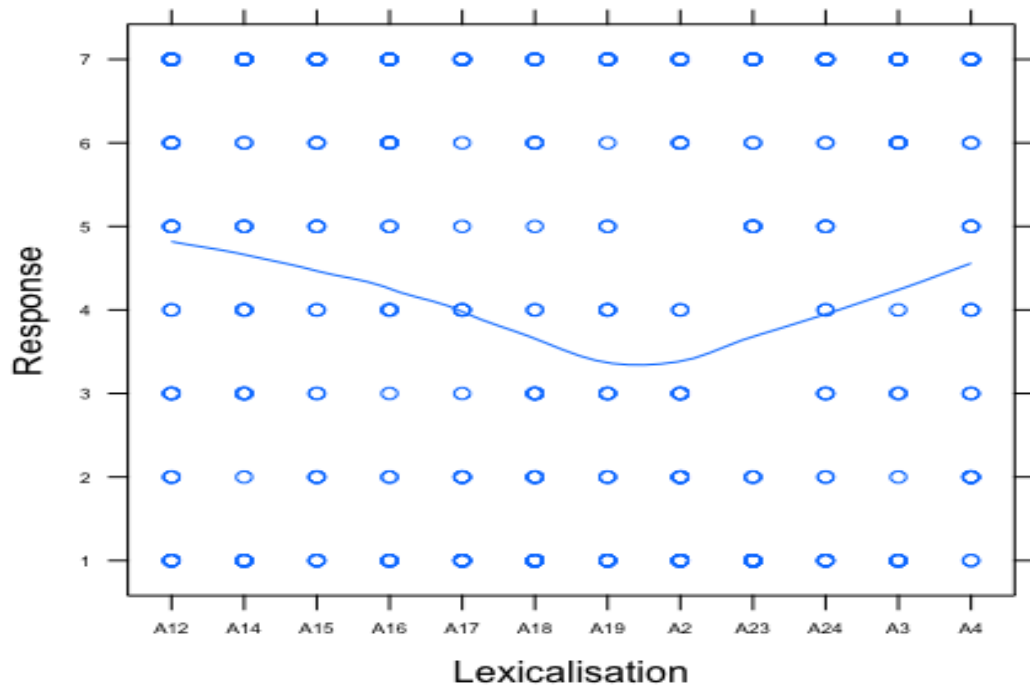


Figure 4-5 Inter-individual variations in the rating of bare *wh*-questions with RPs in the NonCrossed_Island and Crossed-NonIsland conditions.

Ratings are not affected by lexicalisation.

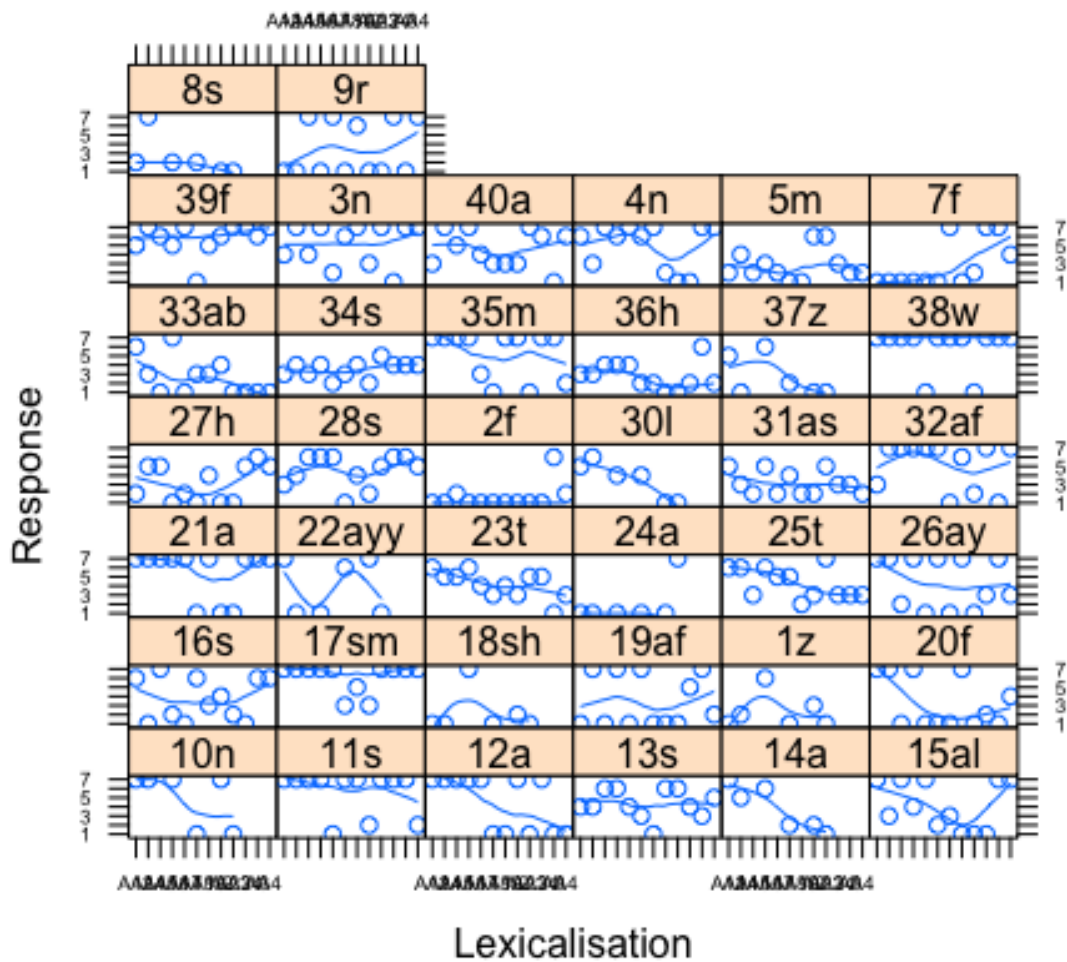


Figure 4-6 Inter-individual variations in the rating of bare *wh*-questions with RPs in the NonCrossed_Island and Crossed-NonIsland conditions.

Ratings are not affected by item by subject effect.

As for the effect of WM capacity, results revealed that it has a limited effect on acceptability. The only significant effect of WM capacity on the acceptability of bare *wh*-questions is that participants with higher WM scores accept RPs in Crossed Non-Island conditions more than participants with lower WM scores (Estimate: -0.32, z: -3.17, p: 0.002).

Moving to *illi*-structures, acceptability is found to be mainly affected by whether an RP or a gap appears at the dependency tail. In other words, there was a main effect of Tail, such that conditions created with RPs were rated significantly higher than gapped conditions. In more detail, Non-Crossed Island conditions: (Estimate: -5.5, z: -15.21, p: < 0.0001); Crossed Non-Island conditions: (Estimate: -5.62, z: -15.8, p: < 0.0001); Crossed Island conditions: (Estimate: -4.03, z: -15.6, p: < 0.0001).

There is a main effect for Condition, such that the Non-Crossed Island condition with RP is rated higher than Crossed Island with RP (Estimate: -2.16, z: -6.48, p: <.0001), and the Crossed Non-Island condition with RP is rated higher than Crossed Island with RP (Estimate: -2.07, z: -6.19, p: <.0001). However, this degradation in acceptability does not alter the overall acceptability rating pattern: i.e. Crossed Island with RP is still rated in the acceptable range (M=2.06, SD=1.98), while Crossed Island with gaps is rated lower than Crossed Island with RP (Estimate: 4.02, z: 15.7, p: < 0.0001). Moreover, the results do not reveal an effect of embedding: there is no significant difference between the rating of Non-Crossed Island condition with RP conditions and Crossed Non-Island condition with RP conditions (Estimate: -0.09, z: -0.21, p: 0.8).

As for the effect of WM capacity, results revealed that it has very little impact on acceptability. Its only significant effect on the acceptability of *illi*-structures is that participants with larger WM scores tend to accept gaps in Crossed-Island conditions more than participants with smaller WM scores (Estimate: -0.27, z: -3.07, p: 0.002).

As stated above, we used different lexicalisation sets across the different conditions in the current study. To ensure that this design did not affect the robustness of the statistical analysis, we plotted the predicted values of different lexicalisation sets across the different conditions in the current study in Figure 4-7. The plots revealed that while there is variation in how lexicalisation affects the rating score, this occurs in a way that is consistent with our design: in all lexicalisations, the relative rating of the conditions is consistent with the results of the model. This confirms the robustness of the findings across lexicalisations.

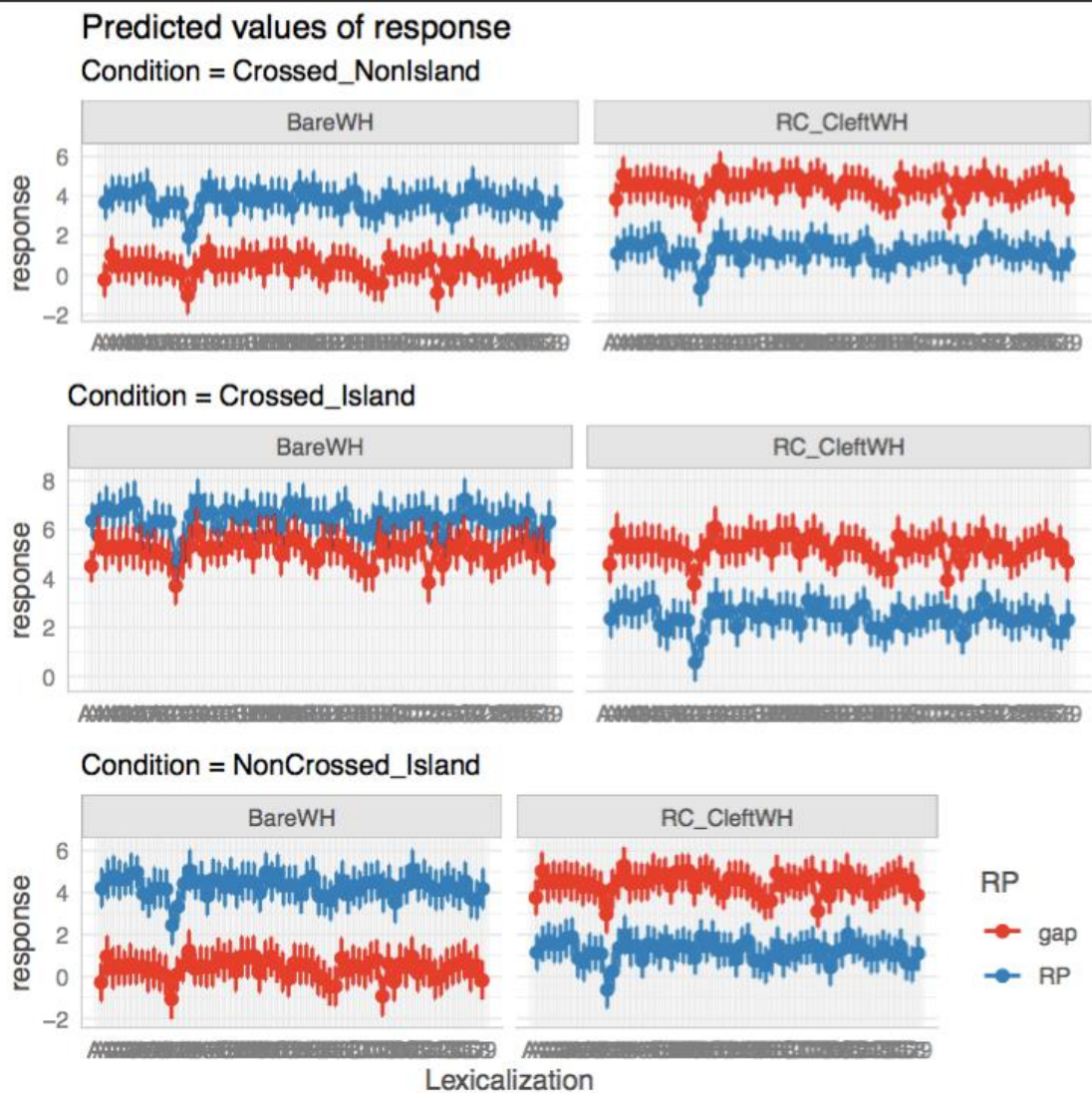


Figure 4-7. Predicted values of lexicalisation across the different conditions in the acceptability judgment study.

4.4.2.4.2 Difference as dependent variable

In this analysis, we calculated the difference in acceptability between each minimal pair of gap dependencies and RP dependencies. The means of these differential scores across subjects were then regressed against the Condition [of the embedded clause] and Structure [of the *wh*-dependency]. The difference score was measured as (Gap [rating] – RP [rating]). Recall that on the acceptability scale, 1 is fully acceptable, and 7 is fully unacceptable; hence, a negative difference score indicates that the sentence with a gap was preferred to its variant with RP, while a positive difference score means that the sentence with RP was preferred to its variant with a gap. This is illustrated more fully in (20).

20. **Example of calculating the difference score:**

If, in a minimal pair, the Gap-sentence (21a) is rated as 2 (acceptable) and the RP-sentence (21b) is rated as 7 (unacceptable), we will obtain a negative difference score, meaning that Gaps are preferred:

21. Minimal Pair:	Rating
a. What is the car that you said that Muhammad hit [gap]?	2
b. What is the car that you said that Muhammad hit it?	7

Difference in rating scores = Gap (rating) – RP (rating) = 2 – 7 = -5

On the other hand, if the RP sentence (21b) is rated as 2 (acceptable) and the Gap sentence (21a) is rated as 7 (unacceptable), we will obtain the following positive difference score, meaning that RPs are preferred:

Difference in rating scores = Gap (rating) – RP (rating) = 7 – 2 = 5

Table 4-7 below illustrates the general descriptive statistics in regard to the means of difference scores across Conditions in each dependency structure.

Table 4-7 Means of difference scores calculated across conditions as (acceptability score of Gap - acceptability score of RP).

Positive scores indicate a preference for RPs; negative scores indicate a preference for gaps.

Descriptive statistics			
Structure	Condition	Difference.mean	Difference.sd
BareWH	NonCrossed_Island	-3.111	2.435
BareWH	Crossed_NonIsland	-2.533	2.649
BareWH	Crossed_Island	-0.265	2.497
<i>illi</i> -structures	NonCrossed_Island	4.205	2.260
<i>illi</i> -structures	Crossed_NonIsland	4.191	2.350
<i>illi</i> -structures	Crossed_Island	3.873	2.824

Consistent with the results of the previous analysis, the means of the difference scores are negative across the different conditions in bare *wh*-questions, suggesting that gaps are always preferred in these structures. In *illi*-structure dependencies, on the other hand, difference scores are always positive, suggesting that RPs are consistently preferred to gaps across the different conditions. The statistical analysis below will reveal whether or not these observations are significant.

Modelling procedures

Difference scores were analysed with Generalised Additive Models (Wood, 2011) using the *mgcv* package in R Studio (Version 1.8.25). The models in this analysis are fitted starting from random effects only, after which fixed effects are added incrementally. The first model was constructed with Subject and Lexicalisation as random factors. The basic model was then extended to include one variable each time; if this improved the model fit, this variable was kept and another one was added. That is, fixed effects were added one by one and retained only if they improved the model's fit, as indicated by *compareML*.

Likelihood ratio comparisons between nested models revealed that the optimal model was the one that included an interaction between Structure (BareWH and *illi*-structures) and Condition (NonCrossed_Island, Crossed_NonIsland, Crossed_Island), a random intercept for Lexicalisation and Subject, and random slopes for Subject by Condition and Subject by Structure.

Results of the optimal model are summarised in Table 4-8 . Coefficients of random effects are presented in Table 4-9.

Table 4-8 Summary of the parametric terms from the Generalised Additive Model fitted to Difference data.

Reference level: NonCrossed_Island_BareWH.

Variables	Estimate	Std.Error	z value	P.value
Intercept	-3.12	0.35	-8.99	<0.0001
Crossed_NonIsland	0.6	0.4	1.44	0.15
Crossed_Island	2.91	0.35	8.27	<0.0001
illi-structures	7.3	0.40	18.15	<0.0001

Crossed_NonIsland: illi-structures	-0.52	0.5	-1.13	0.26
Crossed_Island: illi-structures	-3.13	0.41	-7.7	<0.0001

Table 4-9 Coefficients of random factors

Groups	Edf	p-value
s(subject)	9.115	0.124
s(subject,Condition)	26.21	<0.0001
s(subject,structure)	65.62	<0.0001
s(Lexicalisation)	38.9	<0.0001

The coefficients of the model are plotted below for ease of interpretation.

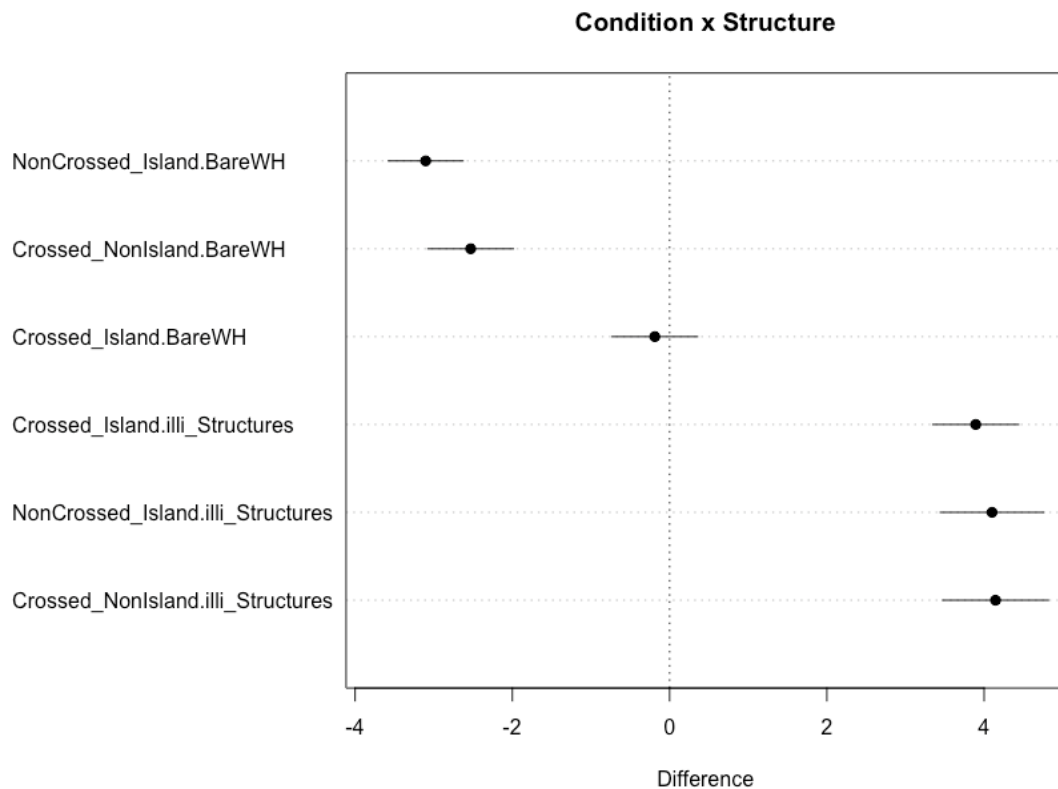


Figure 4-8 Difference scores calculated as (gap – RP) in *illi*-structures and BareWH.

Negative scores indicate that gaps are preferred; positive scores mean that RPs are preferred.

The results revealed a main effect for structure, such that the Difference score is always higher in *illi*-structures in comparison to BareWH structures (Estimate: 7.3, t: 18.15, p: <0.0001), suggesting that RPs are consistently preferred to gaps in *illi*-structures. In bare *wh*-questions, results revealed that the difference between the acceptability of gap and RP dependencies differs significantly between Crossed Island and Non-Crossed Island conditions (Estimate: 2.91, t: 8.27, p: <0.0001), as well as between Crossed Island and Crossed Non-Island conditions (Estimate: 2.34, t: 7.30, p: <0.0001). However, there is no significant difference between Crossed Non-Island and Non-Crossed Island conditions (Estimate: 0.6, t: 1.44, p: 0.15).

In *illi*-structures, there is no significant difference between difference scores across the following conditions: Crossed Island conditions compared to Non-Crossed Island (Estimate: -0.21, t: -0.84, p: 0.4), Crossed Island conditions compared to Crossed Non-Island (Estimate: -0.25, t: -1.02, p: 0.31), and Non-Crossed Island conditions compared to Crossed Non-Island conditions (Estimate: -0.044, t: -0.15, p: 0.87).

4.4.3 Discussion

The present study primarily aims to investigate the types of RPs used in Baha Arabic. This was achieved through an investigation of how crossing non-island and island clause boundaries affect resumption behaviour across different types of *wh*-dependencies in Baha Arabic. We included three types of *wh*-dependencies in the current study: bare *wh*-questions, relative clauses and cleft *wh*-questions. The hypotheses drawn from the relevant literature concerning the distinct behaviour of true and intrusive RPs are listed below:

- Sensitivity to islands is a diagnostic for syntactic movement.
- Features of C affect the type of dependency; this implies that resumption behaviour differs across different types of dependency structures.
- True RPs appear in binding dependencies and fully restore their acceptability.
- Intrusive RPs appear in movement dependency and partially improve acceptability.

- The extent to which resumption can ameliorate islands varies across types of island: the literature indicates that the facilitation effect of RPs is more evident in weaker islands, like adjunct islands, than stronger islands like RC islands (Tucker et al., 2019).
- Individuals with lower WM capacities will be more accepting of intrusive RPs than individuals with higher WM capacities in cases of demanding dependencies (islands and long-distance dependencies).

Starting with *illi*-structures, the results indicate that overt RPs are obligatorily required at the dependency tail; these dependency structures receive high ratings when a RP appears at the dependency tail, regardless of whether or not it appears in a short-distance dependency, and regardless of whether the dependency crosses an island clause or a non-island clause boundary. Essentially, the acceptability of *illi*-structures seems to largely be affected by the presence or absence of resumptive pronouns, such that RP dependencies, regardless of condition, are rated in the acceptable range, while their gapped counterparts received a rating in the unacceptable range. These results suggest that binding relations are involved in the derivation of these dependency structures, as well as that genuine RPs obligatorily appear at the dependency tail as a requirement for establishing the syntactic binding dependency at LF between the filler phrase and the pronominal element. These results are therefore compatible with the hypothesis that the complementiser *illi* initiates a binding dependency that must terminate with a pronominal element at the integration site.

Moving on to the sensitivity of binding dependencies to processing costs, Alexopoulou (2010) developed a hypothesis that built on the assumptions of the Syntactic Prediction Locality Theory (Gibson, 1998) to propose that binding dependencies are insensitive to (i) memory cost (the processing cost that arises due to storing the filler phrase that is to be integrated later into the dependency structure) and (ii) integration cost (the processing cost that arises due to integrating the filler phrase into the dependency). In particular, Alexopoulou (2010) linked integration and memory processing costs to the successive-cyclic nature of movement dependencies and hence argued that the binding resolution of syntactic dependencies does not induce these costs.

However, as we argued earlier by building on assumptions of WM-based theories of processing FGDs (e.g. resource limitation theory) and findings from experimental studies, the online processing of binding and movement FGDs is sensitive to WM capacity-related processing costs. In other words, regardless of the FGD's underlying

syntactic structure, parsers need to keep the filler phrase active in WM throughout the processing of FGDs until the integration site is reached. Maintaining the filler phrase in WM while processing a non-island/island clause boundary is costly in terms of WM resources. This hypothesis therefore predicts that relative clauses and cleft *wh*-questions with RPs would be more costly to process when a non-island clause boundary is crossed (i.e. Crossed Non-Island Condition), as well as when an island clause boundary is crossed (i.e. Crossed Island condition); this, in turn, might translate into lower acceptability ratings.

The results of the current study show that *illi*-structures with RPs are accepted significantly less in the Crossed Island condition compared to other conditions. These findings indicate that *illi*-structures with RPs are sensitive to the processing cost incurred by the crossing of an island boundary. To illustrate, a marginally (but significantly) lower acceptance rate is observed for RP dependencies that cross island boundaries compared to their counterparts where no island boundaries are crossed. However, this decline in acceptability is marginal and has no effect on the overall pattern of acceptability ratings; sentences in the crossed-island condition are still rated as highly acceptable. The marginal lower acceptance of RP dependencies that cross islands in *illi*-structures when compared to their counterparts where no islands are crossed can best be understood as reflecting constraints on processing, rather than syntax. In particular, due to the modest impact of islands on the acceptability of *illi*-structures, this decline in acceptability should be interpreted in terms of processing rather than syntactic constraint violation. Holding the filler phrase in WM while processing an island clause boundary, the semantics of which are inherently complex, overloads the parser's WM resources, leading to a perception of reduced acceptability. However, no such effect was detected across long-distance dependencies that do not cross an island boundary. In short, there is no significant difference between the acceptability of Non-Crossed Island and Crossed Non-Island conditions with RPs.

In summary, the results of the offline experiment indicate that the distributional properties of RPs in *illi*-structures in Baha Arabic are neither conditioned by processing factors (i.e. embedding) nor regulated according to last resort considerations (i.e. used only inside islands); rather, RPs are obligatory in *illi*-structures regardless of islandhood or level of embedding.

Moving on to bare *wh*-questions, we found that the pattern of acceptability is largely affected by whether or not a dependency structure crosses an island clause boundary;

in other words, results indicate that *wh*-dependencies crossing an island clause boundary are rated as unacceptable, regardless of whether a gap or an RP appears at the dependency tail. The observation that bare *wh*-questions, with both gaps and RPs, are sensitive to islands suggests that (i) these dependency structures are derived by movement and that (ii) the RPs in such structures are intrusive, i.e. not grammatically licensed.

These results are compatible with Alexopoulou's (2010) argument that intrusive RPs appear in movement dependencies, and that these RPs are unable to alter the overall pattern of the acceptability of island-violated dependencies. However, no amelioration effect of these intrusive RPs on islands was observed: RPs are perceived to be as bad as gaps inside islands.

When dependencies do not cross an island clause boundary, bare *wh*-questions with gaps are rated as highly acceptable. There is also a significant difference between the acceptability of non-crossed islands and crossed non-island conditions, suggesting that participants were sensitive to long-distance dependencies.

However, the behaviour of RPs in these conditions raises suspicions as to whether RPs are in fact not generated by the grammar in *wh*-questions. In general – unlike gaps, which are rated as acceptable – RPs received inconsistent ratings in non-crossed islands and crossed non-island conditions. In particular, the sum of the ratings received for resumption in these conditions target the middle point of the scale. However, as illustrated above, careful investigation of participants' ratings for these conditions revealed that there was strong variability within participants, such that each participant rated sentences in this condition as sometimes acceptable, sometimes unacceptable, and sometimes in the middle of the scale.

Surprisingly, the distributional properties of RPs in bare *wh*-questions are not determined either by last resort considerations or by processing factors. More specifically, the results revealed that RPs in bare *wh*-questions in Baha Arabic are sometimes accepted in non-island (non-demanding) contexts, while they are generally rejected in English under these circumstances. Furthermore, the results provided a suggestion that higher WM scores were more associated with higher acceptance of RPs in non-crossed island and crossed non-island conditions than lower WM scores. This outcome is not anticipated by processing theories of RPs.

We thus present the following possible explanation for the status of RPs in bare *wh*-questions. RPs in Baha Arabic bare *wh*-questions are as intrusive as RPs in English.

However, since Baha Arabic does allow RPs in certain constructions (such as *illi*-structures), it is possible that the presence of a resumptive structure in the everyday grammar might lead Baha Arabic speakers to be less biased against sentences with RPs more generally when there is no violation of a grammatical constraint. However, if Baha Arabic RPs in bare *wh*-questions are as intrusive as RPs in English, why do these structures not ameliorate island effects? We propose that the absence of their facilitative effect on islands might be due to the following reasons: (i) the use of offline acceptability judgements, (ii) experimental sentences not being contextualised, (iii) Lack of referential contexts that might help parsers to interpret the *wh*-filler phrase referentially. We address these in turn below.

Use of offline acceptability judgements. As illustrated in section 3.4.2, while the facilitation effect of intrusive RPs in English has never been detected in offline acceptability experiments (Heestand et al., 2011; Keffala, 2011; Polinsky et al., 2013), it has been observed in studies utilising experimental methods such as self-paced reading (Hofmeister and Norcliffe, 2013; Hammerly, 2019), comprehensibility judgement (Beltrama and Xiang, 2016), forced-choice tasks (Ackerman et al., 2018) and acceptability judgements performed while engaged in cognitive tasks that constrain WM resources (Chacón, 2019). As noted by Phillips et al. (2019, p.7) in his discussion of the experimental methods used in the investigation of resumption in English, ‘scalar acceptability ratings can be a blunt tool, especially when a rating for an entire sentence is used as a proxy for the status of one specific piece of that sentence, such as a RP’ (7). However, in the current study, acceptability judgement revealed a significant improvement effect in cases of RPs in relative clauses and cleft *wh*-questions. Hence, Phillip’s argument might only apply to intrusive RPs; that is, if island-violated dependencies with RPs are not grammatical as island-violated dependencies with gaps, identifying a difference in their acceptability pattern is not guaranteed.

Lack of contextualisation. The experimental sentences in the present study were all presented without preceding context. As noted by Hofmeister (2012), presenting direct questions in this way might increase the processing cost of such structures; more specifically, presenting direct questions that are not contextualised causes these questions to sound pragmatically odd, which makes them more complex to process. This, in turn, might further hinder the emergence of any facilitation effect of intrusive RPs. In other words, the facilitation effect of RPs might be not detected due to the absence of any context that might increase the plausibility of these questions

and thus indirectly allow for the facilitation effect to be detected. It is therefore possible that presenting experimental sentences along with preceding context will increase the likelihood that the RP will be accepted, which might in turn (i.e. indirectly) allow the facilitation effect to emerge. The lack of such preceding contexts in the current study might thus affect the acceptability patterns of RPs in bare *wh*-questions.

Lack of referential contexts. The facilitation effect of RPs has been attributed to their discourse nature (Ariel, 1999; Frazier and Clifton, 2002; Alexopoulou, 2010; Beltrama and Xiang, 2016; Chacón, 2019). In other words, intrusive RPs facilitate processing by allowing parsers to resolve the syntactic dependency anaphorically. It is therefore possible that the lack of processing effect for RPs is due to the difficulty of establishing an anaphoric dependency between the *wh*-phrase ‘what’ and the pronouns in the absence of referential contexts. More specifically, it is possible that the facilitation effect of RPs will emerge if experimental sentences are presented along with preceding referential contexts, which will allow parsers to interpret the *wh*-filler phrase anaphorically. Beltrama and Xiang (2016) found that presenting experimental sentences along with preceding referential context sentences in English and Italian allows the facilitation effect of RPs to emerge.

That being said, it is unclear, under the working hypothesis that RPs in bare *wh*-questions are intrusive, why participants with larger WM scores would accept RPs in non-island contexts more than participants with smaller WM scores.

An alternative interpretation for the behaviour of resumption in bare *wh*-questions is that participants might have accepted non-island bare *wh*-questions with RPs through accommodation, by topicalising the *wh*-phrase (i.e. interpreting these *wh*-questions as Clitic Left Dislocation (CLLD)). According to Aoun et al. (2009), CLLDed elements in Arabic are unambiguously interpreted as topics, and are obligatorily linked to a pronoun inside the clause. Typical example of this construction in Arabic is given in (22)

22. at-tilmiiḏat-u raʔaa-ha saami l-baariḥa
 the-student.fs-Nom saw.3ms-her Sami the-yesterday
 ‘The student, Sami saw her yesterday.’

(Aoun et al., 2009)

Speakers with higher WM capacity would therefore be able to summon contexts more easily to accommodate resumption (interpret filler phrases as topics). But it remains

unclear, under this interpretation, why *wh*-fillers in island-violating dependencies with RPs are not interpreted as topics, as is the case in non-island dependencies.

In summary, the results of the current study are problematic for theories that argue against the existence of a distinction between true and intrusive RPs. Instead, they are consistent with accounts that propose a typological distinction between true and intrusive RPs. Evidence for island-effects in *wh*-questions where RPs are not allowed, along with the absence of such an effect in relatives and cleft *wh*-questions where RPs are obligatory, support (i) the grammatical accounts' claim that there are two mechanisms for encoding syntactic dependencies (i.e. movement-relations that are sensitive to island-effects (*wh*-questions) and binding-relations that are insensitive to island-effects (relatives and cleft *wh*-questions)), and (ii) the claim that true RPs appear in binding dependencies while intrusive RPs appear in movement dependencies.

Furthermore, results are supportive for the argument that the distinction between binding vs movement dependencies, and henceforth between the two types of RP, is linked to the morpho-syntactic features of elements that appear in the CP layer (Alexopoulou, 2010; McCloskey, 2017); in other words, elements in C determine whether or not a dependency structure is derived by a binding dependency, which must terminate with a pronominal element at the site of integration, or derived by movement relations, which allows intrusive RPs to appear at dependency tail in certain contexts.

The impact of processing-related factors on the processing of FGDs is furthermore evident in the current study: (i) long-distance dependencies with gaps are rated significantly lower than short-distance dependencies with gaps in bare *wh*-questions (embedding effect), and (ii) dependencies crossing islands in *illi*-structures with RPs are rated significantly lower than dependencies that do not cross islands in *illi*-structures with RPs (although both are grammatical).

Furthermore, results revealed that the type of island affects resumption across the different types of *wh*-dependencies similarly.

4.4.4 Summary of results

The hypotheses in 4.3 are derived from the assumption that resumptive pronouns that appear in binding dependencies are true RPs, while RPs that appear in movement dependencies are intrusive RPs. This distinction between these two types of

dependencies is linked to the features of C. When featural specifications of C trigger movement, intrusive RPs are used; when featural specifications of C do not trigger movement, true RPs are obligatorily inserted so that a binding dependency is established between C and the pronoun at LF.

Of interest to us here is the Arabic relative complementiser *illi*. Alexopoulou (2010) argued that this complementiser belongs to the class of complementisers that do not trigger movement, and hence that true RPs are obligatory with *illi*. In the present study, we tested two types of *wh*-dependencies where the complementiser *illi* is obligatorily used: relative clauses and cleft *wh*-questions. Hence, if this is the type of complementiser that affects the type of dependency, and accordingly the type of RPs, we predict that relative clauses and cleft *wh*-questions should be derived by binding relations and feature obligatory true RPs. These dependency structures are insensitive to island effects. In bare *wh*-questions, on the other hand, operators in C trigger movement and hence should be sensitive to islands, even when intrusive RPs are used. Hence, the results of experiment 1 support the hypothesis that cleft *wh*-questions and relative clauses, which share the obligatory appearance of the complementiser *illi*, are derived by binding relations, while bare *wh*-questions are derived by movement.

In summary, the results of the current study on Baha Arabic have revealed the following:

- There are two types of dependency structures in Baha Arabic:
 - Binding dependencies in relative clauses and cleft *wh*-questions:
 - RPs are quasi-obligatory in relative clauses and cleft *wh*-questions
 - Relative clauses and cleft *wh*-questions with RPs are marginally less accepted when crossing islands
 - Both type of islands, adjunct and RC-islands, impact acceptability similarly
 - Distance has relatively little impact on acceptability
 - Movement dependency in bare *wh*-questions:
 - RPs are never fully accepted in bare *wh*- questions (only marginally so, in non-islands)

- bare *wh*-questions are strongly rejected when crossing islands (whether or not they feature a RP)
- Both type of islands, adjunct and RC-islands, affect acceptability similarly
- Distance has a modest but significant impact on acceptability

4.4.5 Limitations and questions

This experimental study has some potential limitations. First, experimental sentences were not followed by comprehension questions; hence, it is possible that participants did not really attempt to understand the structures. Second, the experiment was too long; as participants who completed the two surveys read a total 412 sentences (1/3 of which were experimental sentences), it is possible that the results are affected by fatigue effects. Third, and most importantly, offline methods are not optimally suited to investigating the processing effect of intrusive RPs. Since independent effects of resumption in the context of extraction have been confirmed using processing measures, and since intrusive resumptive pronouns occur more often in produced speech, as in English (Kroch 1981; Prince 1990), it is possible that in real-time applications where temporal and memory constraints are involved, the lack of isolated effects of resumption on acceptability ratings in the present study may be an artifact of the use of an offline measure; that is, the effect of working memory capacity may be reduced in such experiments, as there are no temporal constraints and participants are potentially able to reread sentences several times before making a decision.

It is therefore important to move from investigating the assumptions of processing theories in offline experiments, where participants are consciously making decisions about the acceptability of sentences, to online experiments, where participants' unconscious reactions to the stimuli can be measured.

Additionally, the source of the decline in acceptability for the Crossed Island condition in relatives and cleft *wh*-questions is still not clearly understood from the acceptability judgement study. The extent of the processing effort necessary to resolve these dependencies can only be investigated via online methods (e.g. self-paced reading registering reaction times).

Moreover, it is possible that presenting direct questions that are not contextualised causes bare *wh*-questions to sound pragmatically odd, which makes them more

complex to process, and accordingly, hinder the emergence of any processing effect that intrusive RPs might have.

Finally, it is possible that the specific pragmatic features of the filler phrase may have an effect on the acceptability of resumption in *wh*-questions. In other words, the lack of amelioration effect of RPs in bare *wh*-question might be due to the lack of referential contexts that might help parsers to interpret the *wh*-filler phrase referentially.

4.4.6 Outstanding questions

1. Would online measures reveal a different picture of acceptability judgements?

Specifically:

- What is the processing cost associated with long-distance dependencies?
- What is the processing cost associated with crossing an island boundary?
- Do RPs alleviate the processing cost of *wh*-questions?
- Do RPs affect real time processing differently across types of *wh*-dependencies?

2. What are the discourse properties of RPs in *wh*-questions? Are they sensitive to discourse licensing?

The first question will be addressed in Chapter 5, while the second question will be investigated in Chapter 6.

4.5 Conclusion

The data presented here revealed that Baha Arabic features grammatical (true) RPs in relative clauses and cleft *wh*-questions. The results indicate that RPs appearing at the tail of binding dependencies in cases of relative clauses and cleft *wh*-questions are genuine RPs, a conclusion supported by the high acceptance rate when RPs are featured.

The grammatical status of RPs in *wh*-questions is not yet clear. RPs are never fully accepted in bare *wh*-questions (only marginally so, in non-islands). This pattern of acceptability suggests that these pronouns are intrusive like RPs in English. However, results revealed that RPs in bare *wh*-questions are sometimes accepted in non-island (non-demanding) contexts, while they are generally rejected in English. We speculate that, since Baha Arabic does allow RPs in constructions such as *illi*-structures, it is

possible that the presence of a resumptive structure in the grammar might lead Baha Arabic speakers to be less biased against sentences with RPs generally, even when no violation of a grammatical constraint occurs. Alternatively, it is possible that the issue here is not intrusive vs true resumption but CLLD-style, i.e. the possibility to topicalise the *wh*-phrase; parsers accepted bare *wh*-questions with RPs when they interpret *what*-fillers as topics (as in CLLD structures).

Furthermore, the results do not reveal any amelioration effect of RPs inside islands in bare *wh*-questions. We argued that this might be an artifact of using an offline acceptability judgement experimental method. It is possible that acceptability judgement experiments are not ideal for investigating the role of RPs in FGDs in instances where RPs are not grammatically licensed. To further investigate the role of RPs in the comprehension of FGDs in Baha Arabic, we will use online measures in the next chapter.

As for the source of island effects, these results support the grammatical theory of islands, which draws a distinction between binding and movement dependencies. In particular, the different acceptability patterns of island-violating dependencies in relatives and cleft *wh*-questions on one hand and bare *wh*-questions on the other hand suggest that the two types of dependencies do indeed have different underlying syntactic structures that determine the status of islands. However, despite the fact that RPs are grammatically licensed inside islands in relative clauses and cleft *wh*-questions, crossing an island clause boundary results in a small (but statistically significant) decrease in acceptability. This decline in acceptability most likely reflects the complexity associated with processing island-violating dependencies, which is in accordance with WM-based accounts of island effects.

Chapter 5 Processing RPs in Baha Arabic *Wh*-dependencies

5.1 Introduction

Having established in the previous chapter that Baha Arabic features different types of RPs across *wh*-dependencies, the present study focuses on *wh*-questions, with the aim of assessing the processing role of RPs via a self-paced reading task.

The acceptability judgement experiment conducted in the previous chapter revealed that *wh*-questions and *illi*-structures behave differently in terms of their interaction with islandhood and resumption, which in turn suggests that two different types of dependencies are involved in their derivation. We concluded that *illi*-structures, where RPs are obligatory regardless of islandhood, are derived by binding relations, as the presence of a dependency that crosses an island boundary does not alter the overall acceptability of these structures. On the other hand, movement relations are involved in the derivation of *wh*-questions, as evidenced by their sensitivity to island effects.

Consequently, we hypothesised that *illi*-structures involve true RPs, and that the syntactic binding dependency is established between the filler phrase and the pronoun at LF (McCloskey, 2002; Alexopoulou, 2010). Bare *wh*-questions, on the other hand, are movement dependencies where intrusive RPs can be used to facilitate processing. However, the processing effect of RPs could not be detected in bare *wh*-questions in the offline acceptability judgement study.

Furthermore, the results of the acceptability judgement experiment revealed that, despite the observation that island-violating *illi*-structures with RPs are rated as acceptable, there is still a slight (but significant) decrease in their acceptability compared to their non-island counterparts. We interpret this effect as the processing cost of crossing an island boundary. In the current study, the difference between *illi*-structures that cross islands vs non-islands will be used as a baseline for determining the processing cost of crossing an island boundary. This will allow us to disentangle the effect of syntactic violations from the effect of processing cost in *wh*-questions. Consequently, in this chapter, we aim to assess whether RPs facilitate the processing of bare *wh*-questions in Baha Arabic, using relative clauses as a benchmark for the processing cost of islands.

The remainder of this chapter is organised as follows. Section 5.1.1 reviews the processing theories of island effects in binding and movement dependencies. Section 5.1.2 reviews the processing function of intrusive RPs. We will next review the impact

of island type on the processing advantage of intrusive RPs in 5.1.3, while the relation between the processing advantage of RPs and individual differences in WM capacities is investigated in 5.1.4. A summary of the research questions and predictions will be presented in 5.1.5. Next, Section 5.2 presents the methods of Study 2, summarises its findings and discusses its limitations. Section 5.3 then presents Study 3, in which we test the impact of contextualising experimental sentences, as well as manipulating the d-linking properties of fillers, on the processing of resumption in Baha Arabic *wh*-questions. Section 5.3.6 summarises findings of Study 3, after which section 5.3.7 presents the outstanding questions. Finally, Section 5.4 concludes the chapter.

5.1.1 Processing costs of islands

The current study decomposes the island effect into two components: the processing cost of islands independent of movement (in *illi*-structures) and the additional cost of a movement dependency across an island boundary (in bare *wh*-questions).

As a binding dependency, *illi*-structures are not affected by grammatical constraints on movement, as suggested by the results of the acceptability judgement experiments; that is, they are not prone to island violations. However, *illi*-structures are still sensitive to the cost of processing island-violating dependencies. More specifically, the processing of any Filler Gap Dependency (FGD) requires holding the filler active in WM until the integration site is reached. This applies to binding and movement dependencies alike.

For instance, according to the resource limitation theory (RLT) (Kluender, 1991; Kluender and Kutas, 1993b) the cost of holding the filler phrase active in WM while processing clause boundaries depends on the degree of semantic complexity associated with that clause boundary. This cost is particularly high when the crossed clause boundary is an island. As WM capacity is limited, filler phrases in WM are vulnerable to decay effects when processing demands are high.

Since *illi*-structures are binding dependencies, but still show some level of sensitivity to island effects, we proposed that this decline in the acceptability of island-violating *illi*-structures with RPs when compared to non-island conditions reflects the processing cost of crossing an island clause boundary while the dependency remains unresolved. Following Michel (2014), we hypothesise that the processing cost (i.e. the maintenance cost) will increase at the island boundary, as it represents the combined

processing costs of maintaining a filler phrase in WM while crossing a complex clause boundary. We anticipate that this effect will be observed in both *illi*-structures (binding) and bare *wh*-questions (movement).

The processing cost of islands can also be observed at the integration site. When the level of activation of a filler phrase is lowered in WM due to the presence of an island clause boundary, the filler phrase becomes inaccessible for either retrieval or reactivation at the integration site, as the retrieval cost is too high (Hofmeister and Sag, 2010). The retrieval cost at the integration site then spills over onto subsequent word regions (Lewis and Vasishth, 2005; Hofmeister, 2011).

The active filler strategy (Frazier and d'Arcais, 1989; Clifton and Frazier, 1989), introduced in Chapter 3, assumes that parsers construct FGDs actively in real time. Experimental investigation further revealed that the active-filling strategy is interrupted inside island structures in movement dependencies, but not in binding dependencies (Keshev and Meltzer-Asscher, 2017). We suggest that this might be due to a WM effect: in short, the cost of reactivating fillers at the integration site is too high for the resolution to occur. This cost is prohibitive in bare *wh*-questions (unless the activation of the filler phrase is boosted), but not in *illi*-structures (which do not feature movement).

Several studies identified a slowdown in RT inside islands in movement dependencies (Hammarly, 2019). This slowdown at the integration site and spillover regions in movement dependencies has been interpreted as an effect of 'the confusion caused by the presence of an unfilled argument position, with no immediately available filler to form a dependency into this site' (Hammarly, 2019, pp.9–10). This implies that the filler phrase is not reactivated from WM in these structures, due to high processing demands as well as the violation of grammatical constraints on movement. Binding dependencies (i.e. *illi*-structures), on the other hand, lack any violation of grammatical constraints as well as the complexity associated with processing movement dependencies; thus, any slowdown in RTs inside islands compared to non-islands in *illi*-structures would simply reflect the processing cost of reactivating fillers in WM.

To summarise, the processing cost of island-violating long-distance dependencies is expected to manifest itself as longer RTs at the clause boundary, the integration site and the spillover region.

1. At the clause boundary:

- (i) Higher cost in islands (compared with non-islands) due to the need to hold the filler in WM while the island is being processed, both in *wh*-questions and *illi*-structures.
2. At the integration site and spillover region:
- (i) Higher cost in island-crossing *wh*-questions due to the syntactic violation compared to island-crossing *illi*-structures.
 - (ii) Higher cost in islands (compared with non-islands) due to the need to reactivate fillers at integration sites, both in *wh*-questions and *illi*-structures.

5.1.2 Processing function of resumption

As we saw in Chapter 3, it was widely accepted that the distributional properties of RPs across languages are modulated by processing factors (Erteschik-Shir, 1992; Ariel, 1999; Hofmeister and Norcliffe, 2013; McCloskey, 2017). This characterisation is proposed for true and intrusive RPs alike. Starting with intrusive RPs, theoretical and experimental investigations indicate that intrusive RPs are used in positions where parsers are not able to hold the filler active in WM due to an increased processing load: for example, in deeply embedded positions (1) and inside island structures (2). In other words, when parsers do not successfully maintain the representation of the filler phrase in WM, or when the filler phrase has a lower level of accessibility in WM by the time gap is processed, RPs can be utilised to help parsers reactivate/retrieve filler phrases in WM (Erteschik-Shir, 1992; Asudeh, 2012; Hammerly, 2019; Chacón, 2019). Furthermore, Alexopoulou (2010) argued that the anaphoric resolution of FGDs initiated by intrusive RPs cancels the complexity associated with the ‘backward’ cyclic resolution of the dependency that requires parsers to integrate traces of the filler phrases at every intervening clause boundary.

1.
 - a. This is the girl that John likes t/*her
 - b. This is the girl that Peter said that John likes t/??her
 - c. This is the girl that Peter said that John thinks that Bob likes t/?her
 - d. ?This is the girl that Peter said that John thinks that yesterday his mother had given some cakes to ?t/her

(Erteschik-Shir, 1992, p.89)

2. a. *I just saw a girl who Long John's claim that ____ was a Venusian made all the headlines.
 - b. I just saw a girl who Long John's claim that she was a Venusian made all the headlines.
- (Ross, 1986, p.260)

As for true RPs, corpus-based studies have revealed that the use of true RPs increases as processing cost increases in both Hebrew (Ariel, 1999) and Irish (McCloskey, 2017). In other words, when languages allow both strategies (the gap strategy and the RP strategy) for establishing FGDs, parsers preferred to use the RP strategy in positions that induce high processing cost; such positions include deeply embedded positions and islands. Alexopoulou (2010) attributed the processing advantage of RP dependencies relative to gap dependencies to the fact that binding dependencies lack the 'backward and forward' cyclic resolution of the dependency that requires parsers to integrate traces of the filler phrases at every intervening clause boundary.

However, the literature also suggests that obligatory RPs, as observed in Baha Arabic *illi*-structures, behave differently from optional RPs with regard to their processing functions. Specifically, we presented in Chapter 1 the findings from Friedmann and Costa's (2011) investigation of the comprehension of resumptive pronouns on two groups of hearing-impaired children and adolescents, who were Hebrew (a language with optional RPs) and Palestinian Arabic speakers (a language with obligatory RPs). Despite the fact that both languages feature grammatical RPs, results revealed a difference in the reaction to RP dependencies between the two groups of participants. In particular, in a picture-matching task, Hebrew-speaking hearing-impaired participants were found to comprehend object relatives with RPs significantly better than their gapped counterparts. This is consistent with the hypothesis that resumption is preferred under performance pressures. Palestinian Arabic-speaking hearing-impaired participants, on the other hand, were shown not to benefit from the presence of resumptive pronouns, as their comprehension of relative clauses in general was always poor. Friedmann and Costa (2011) argued that this finding was due to differences in the nature of resumptives in Palestinian Arabic and Hebrew. Resumptive pronouns in Palestinian Arabic are clitics that need to be licensed by a functional head; they enter the derivation pre-syntactically as part of the numeration. Resumptive pronouns in Hebrew, on the other hand, are full pronouns that enter the

derivation post-syntactically to facilitate the establishment of the dependency in relative clauses.

Our first study on Baha Arabic revealed that RPs do not alleviate island effects in bare *wh*-questions. However, the variability in the acceptance of RPs in bare *wh*-questions suggests that RPs might still have a processing effect. Indeed, we observed a marginal reduction in the acceptability of long-distance *wh*-questions (compared with short-distance ones), along with a marginally higher level of acceptance of RPs in long-distance *wh*-questions. However, the nature of that effect could not be investigated through an acceptability judgement task.

Offline acceptability judgement tasks are not designed to detect differences in processing costs between short and long-distance dependencies (including islands). Indeed, as participants in these conditions can potentially re-read the filler to facilitate interpretation, it is possible that the offline presentation of stimuli would remove the need for RPs at the integration sites; this, in turn, may not allow the facilitation effect of RPs to be observed. Accordingly, the present study will use a self-paced reading paradigm to investigate the role of intrusive RPs in comprehension. If intrusive RPs do indeed facilitate processing of complex FGDs – and, specifically, if they aid in the reactivation of the lost filler phrase in WM – then their occurrence is predicted to facilitate processing in situations where processing demands on WM resources are high. That is, intrusive RPs are expected to speed up RTs as the processing complexity of bare *wh*-questions increases.

As for true RPs in *illi*-structures, since they are obligatory, the question as to whether they facilitate processing in island structures is void. This is because their appearance is required by the grammar. However, we also include *illi*-structures with the gap vs RP manipulation in this study, since one of our aims is to examine the RT correlates of grammaticality judgements.

Furthermore, *illi*-structures (with both island and non-island manipulation) will be used as a benchmark for the processing cost induced by islands in the absence of syntactic movement. We expect that the marginal reduction in the acceptability of islands observed in *illi*-structures in Study 1 will translate into an increase in RT at the clause boundary, integration site and spillover regions (in islands compared to non-islands), reflecting this processing cost.

5.1.3 Types of island

As explained in the previous chapter, the amelioration effect of RPs is affected by the strength of the type of island structure where they occur. As in the previous experimental study, the islands we include are Adjunct and RC islands, both of which are traditionally considered strong islands (Szabolcsi and Lohndal, 2017). However, the status of adjunct islands as strong islands has also been questioned (Cinque, 1990; Truswell, 2007; Farra, 2019; Chaves and Putnam, 2020). For instance, Chaves and Putnam (2020) found that adjunct islands, unlike ungrammatical filler sentences, can satiate over repeated exposures; such a finding is not expected if adjunct islands are indeed strong islands (i.e. a purely syntactic phenomenon).

Furthermore, the literature on processing island-violating dependencies suggests that these two types of island might differ in terms of their processing costs. Keshev and Meltzer-Asscher (2019) argued that *wh*-islands (and possibly RC islands) induce a processing complexity cost that is neither due to dependency length nor to the presence of an island structure. Specifically, in *wh*-island (and RC island) structures, the embedded island structure introduces a second filler that needs to be actively maintained in working memory. In addition, in many cases, the two maintained dependencies occur in adjacent positions, requiring the near-simultaneous retrieval of both fillers. The second filler might constitute a prominent distractor and thus interfere with the retrieval of the first filler. In addition, RC islands, unlike adjunct islands, are a kind of 'hybrid' dependency (binding inside the RC, movement outside it). Adjunct islands, on the other hand, are not subject to these processing costs.

Literature on the processing advantage of RPs suggests that intrusive RPs improve the acceptability of weaker islands more than stronger islands (Alexopoulou and Keller, 2007). For instance, Tucker et al.'s (2019) experimental investigation of RPs in MSA revealed that the strongest facilitation effect of RPs is found when RPs occur inside adjunct islands; by contrast, no facilitation effect was found in complex NP islands (or RC island). Tucker et al. (2019) attributed the lack of amelioration effect in CNPC islands to 'the unacceptability of any filler-gap dependency with CNPC constructions more generally' (2019, p.37). We here propose that the observed amelioration effect in adjunct islands is due to the weaker nature of the violation of adjunct islands compared to RC islands.

5.1.4 Cognitive measures of WM capacity and speed of processing

WM and processing speed measures will be also included in the present study to investigate the extent to which individual differences in WM capacities and processing speed affect the processing of RP dependencies. In other words, the inclusion of WM and processing speed measures in our experiments allows us to investigate the extent to which differences in RTs can be explained by differences in parsers' cognitive scores of WM capacities and speed of processing. We further predict a correlation between the cognitive measures of WM capacities and speed of processing.

In addition, we predict that the sensitivity of intrusive RPs to processing factors will be stronger in participants with smaller WM capacities and/or slower processing speed than among participants with larger WM capacities and/or faster processing speed. No such predictions are made concerning the interaction between processing of true RP dependencies and individuals' differences in WM resources or their speed of processing.

5.1.5 Summary of research questions and predictions

5.1.5.1 Research questions

1. Do RPs facilitate processing in *wh*-questions? Specifically,
 - a. Do RPs facilitate processing in more complex host structures (long-distance dependencies and islands)?
 - b. Does the effect of RP vary across island structures?
 - c. Does WM capacity modulate the size of the effect induced by RPs?
2. Will acceptability judgement contrasts translate into RT contrasts in *illi*-structures? Specifically,
 - a. How will the impact of the gap vs RP manipulation and the island vs non-island manipulation in *illi*-structures be reflected in RT data?

5.1.5.2 Predictions

Based on the above discussion, we assume that:

- 1) A facilitation in processing induced by RPs will manifest itself in the form of faster RTs in the spillover regions compared to gap dependencies.
- 2) A disruption in processing induced by RPs will manifest itself in the form of slower RTs in the spillover regions compared to gap dependencies.

Our predictions are listed below:

- 1) RPs will generally be disruptive in bare *wh*-questions but not in RCs.
- 2) In *wh*-questions, the magnitude of the facilitation/disruption effects will depend on the complexity of the host structure (with greater disruption in less complex structures, and/or greater facilitation in more complex structures). We assume the following complexity hierarchy:

Non-crossed island (short-distance) dependency < Non-island (long-distance) dependency < Island (long-distance) dependency.

The facilitation effect of RPs in island dependencies will be more evident in Adjunct island (weaker island) compared to RC-island (stronger island).

- 3) Participants with smaller WM capacities and/or slower processing speed will benefit from a greater processing advantage conferred by RPs than participants with larger WM capacities and/or faster processing speed.
- 4) A disruption in processing induced by crossing islands will manifest itself in the form of slower RTs at clause boundaries and in the spillover regions in island-violating dependencies compared to non-island dependencies.

5.2 Study 2: Processing *wh*-dependencies with resumption

In this section, we present a detailed description of the design of the current experiment. In 5.2.1, we outline Sprouse's superadditivity paradigm, which we had initially intended to use in this experiment. After observing that adopting this design would result in confounding effects, which are discussed in detail in 5.2.1.1, we decided to adopt a partial replication of Sprouse's superadditivity design. This partial design, introduced in 5.2.2, includes three graded conditions, which allowed us to address our research questions while avoiding confounding effects.

5.2.1 The superadditivity paradigm

The current study utilised a self-paced reading experimental method. To address the above research questions, we initially adopted Sprouse's superadditivity paradigm (first introduced in Chapter 3). This paradigm aims to quantify the effects of the extra-grammatical components of island-violating sentences (which involve processing complexity) and thereby separate the effect of the grammatical constraints from the effect of processing constraints. Processing cost is operationalised as (i) the cost of crossing a clause boundary and (ii) the cost of processing an island structure while holding a filler active in WM. Two factors are manipulated, each with two levels: structure (island – non-island) and distance between the displaced phrase and its trace (matrix – embedded). We thus obtain four sentence types (or conditions), as shown in 3:

3. a. Who ___ thinks [that John bought a car]? NON-ISLAND | MATRIX
- b. What do you think [that John bought ___]? NON-ISLAND | EMBEDDED
- c. Who ___ wonders [whether John bought a car]? ISLAND | MATRIX
- d. What do you wonder [whether John bought ___]? ISLAND | EMBEDDED
(Sprouse and Hornstein, 2013, p.314)

The FGD is either short, as in (3a) and (3c), where gaps occur in the matrix clause, or long, as in (3b) and (3d), where gaps occur in the embedded clause. The island structure was either absent (as in (3a) and (3b)), present and not crossed by the *wh*-dependency (as in (3c)), or crossed (as in (3d)).

In essence, this factorial design investigates the possibility that the decreased acceptability of island-violating dependencies is a result of increased processing cost rather than the violation of a grammatical constraint. In Sprouse and Hornstein (2013), the effect of dependency length is calculated as the difference in acceptability between matrix and embedded FGDs, i.e. (3a) vs (3b).

A further decrease in acceptability might arise from the inherent semantic complexity of embedded island structures (as in (3c)) compared to non-island embedded structure (as in (3a)). The structure effect is captured by calculating the difference between the acceptability means of (3a) and (3c).

5.2.1.1 Issues in the design of linguistic stimuli

In Study 1, we controlled the grammatical role and animacy of filler phrases. As explained in section 4.4.1.1, all fillers were required to be objects, as it is impossible to have an overt resumptive in the highest subject position in true resumption languages. All fillers were inanimate to maximise their distinctiveness from the subjects (which were all animate). This also ensured that any facilitation effect of resumption would be independent from the semantic features of filler phrases, as animacy has been argued to improve the acceptability of RPs in Arabic (Aoun et al., 2009) and Greek (Tsimpli, 2003).

Looking back at Sprouse's factorial design sentences, however, we found that animacy and the grammatical roles of filler phrases constituted confounding factors. More specifically, the matrix-dependencies are created with animate subject fillers (4), while embedded dependencies are created with inanimate object fillers (5).

4. a. Who ___ thinks [that John bought a car]? NON-ISLAND | MATRIX
 - b. Who ___ wonders [whether John bought a car]? ISLAND | MATRIX
- (Sprouse and Hornstein, 2013, p.314)
5. a. What do you think [that John bought ___]? NON-ISLAND |
 EMBEDDED
 - b. What do you wonder [whether John bought ___]? ISLAND |
 EMBEDDED

(Sprouse and Hornstein, 2013, p.314)

In Sprouse's superadditivity paradigm, a comparison of the independent processing costs of islands in the matrix condition (where the filler is subject and animate) with the independent processing cost of islands in the embedded condition (where the filler is object and inanimate) may be confounded by the different features of the fillers in these two conditions. In particular, animacy and syntactic roles might have an effect on processing FGDs, as subject dependencies have been argued to be easier to process than object dependencies (King and Just, 1991; King and Kutas, 1995; Polinsky, 2011), while animate fillers were found to have a facilitation effect on the acceptability of island-violating dependencies in Greek and English (Alexopoulou and Keller, 2013). It is therefore predicted that the independent penalty of processing islands in the embedded condition will exceed the independent penalty of processing

islands in the matrix condition due to certain factors (animacy and syntactic role) not being controlled for in the design. That is, the superadditive effect may arise due to the effect of these uncontrolled factors.

We thus attempted to use Sprouse' superadditive design after modifying it in such a way that all experimental sentences were created with inanimate object fillers. However, it subsequently emerged that the baseline condition of Sprouse's superadditivity paradigm (Matrix/Non-island) could not be created with inanimate object fillers under these conditions (as in Study 1), as such sentences are impossible in Baha Arabic. Accordingly, we used adjunct chains rather than object chains in that condition (see (6)). To test the possibility of any potential confounds arising due to chain type, two additional (matrix/island) conditions were created: one with adjunct chain and the other with object chain (see (7), (8)).

6. [Short – No-Island – Adjunct chain]

ʕal ʔayš ʔʕlan-ti ʔnnu ma-raḥ yukuun fiih
On what announced-2SF that no-will be there
muʔtamar laḡawyyat hathi as-sanah?
conference linguistics this the-year?

On what did you announce that there will be no linguistics conference this year?

7. [Short – Island – Adjunct chain]

ʕla ʔayš ʔaḥ mḥammad lamman ruḥ-tum
On what fell-down.3SM Mhammad when went-2PL
al-ḡabah?
the-forest?

On what did Muhammad fall down when you were in the forest?

8. [Short–Island–Object chain]

ʔayš amal zar-at-/=uh lamman kan-at fi
What amal visit-3SF-/it when was-3SF in
ar-ryaḍ?
the-Riyadh?

What did you visit when you were in Riyadh?

However, we then realised that there was indeed a confounding factor induced by these sentences after running the study. Specifically, the problem with sentences created with adjunct chains was that in Baha Arabic, pied-piping is obligatory with gap conditions, while stranding is obligatory with RP conditions (see (9), (10)). This asymmetrical behaviour suggests that, in these particular structures, resumptive pronouns and gaps shouldn't be treated as the same thing.

Pied-piping

9. ʕla ʔayš ʔʕlanti _/(*h) innuh ma raḥ yukun
 On what announce.2F _/(*it) that will-not be
 fiih muʔtamer laḡawyyat haḡi assanah?
 there conference linguistics this year

Stranding

10. ʔayš ʔʕlanti ʕlay * _/(h) innuh ma raḥ yukun
 What announce.2F on* _/(it) that not will be
 fiih muʔtamer laḡawyyat haḡi assanah?
 there conference linguistics this year

'On what did you announce [integration site] that there will be no linguistics conference this year?'

Furthermore, the short non-island condition of Sprouse's superadditivity paradigm cannot be created in Baha Arabic relative clauses; in other words, such sentences cannot be bi-clausal where the second clause is a non-island structure. See (11).

11. Šeft as-sayyarh illi kaled aštra-ha men Jeddah
 Saw.1.P the-car that kaled bought-it from Jeddah
 al-ʔsbuuʕ al-madi.
 the-week the-last

'I saw the car that Kaled bought from Jeddah the last week.'

Accordingly, these sentences were ultimately excluded from the analysis.

5.2.2 Revised design

In light of the above, we decided to forgo the full replication of Sprouse's superadditivity paradigm, and instead implement only the exact same four conditions of the acceptability judgement study (see (12)). While the adjunct-chain conditions in *wh*-questions were included in the experiment, we opted not to include them in the analysis because of the confounding factor induced by the use of adjunct chains. The short non-island relative clauses were further excluded.

12. - Non-Crossed Island / (RP/GAP)
- Crossed Non-Island / (RP/GAP)
- Crossed RC-Island / (RP/GAP)
- Crossed Adjunct-Island / (RP/GAP)

This design is a partial replication of Sprouse's superadditivity design. Although we cannot test for the superadditivity effect of island-violating dependencies, we can still assume a gradient in complexity across conditions. Specifically, the crossed island conditions are expected to be processed slower than the crossed non-island condition, which is in turn expected to be processed slower than the non-crossed island condition. As for the impact of resumption on processing, we predict that the processing advantage of RPs in bare *wh*-questions will be sensitive to the degree of processing complexity across conditions.

We manipulated these conditions in (i) bare *wh*-questions and (ii) relative clauses. We also excluded cleft *wh*-questions, as the results of experiment 1 suggested that they behave similarly to relative clauses. We used the same experimental sentences as in the acceptability judgement study. A summary of the properties of the design of these experimental materials is presented below:

- Sentences created with inanimate, object fillers
- All experimental sentences were bi-clausal and started with a fronted *wh*-phrase. See the examples below for bare *wh*-questions (13) and relative clauses (14); here, the filler phrase is bolded.

13. **ʔayš** amal zar-at/h lamman kan-at fi
What amal visit-3SF-/it when was-2F in
ar-ryaḍ?

the-Riyadh?

'What did Amal visit (it) when she was in Riyadh?'

14. ḥabit **aš-šūqah** illi ʔstʔjar-ti-/ha lamman ruḥ-ti Turkia
 liked-1 the-flat that rented-2F -/it when went-2F Turkey

'I liked the flat that you rented when you went to Turkey.'

Experimental sentences are not lexically matched except for RP and gap dependencies. This was done deliberately to limit the likelihood of a priming effect. However, because these stimuli were intended for use in online measures (self-paced reading), some aspects of the experimental sentences were modified. First, we added a spillover region consisting of a three-word adjunct phrase (except for the non-crossed island condition, in which the integration site is already followed by a spillover region). These adjunct phrases at spillover regions were controlled for length and structure. More specifically, all three-word adjunct phrases consist of a Prepositional Phrase (i.e. a preposition + DP). See below for a sample of the experimental sentences (the bolded parts represent our region of interest). Full materials, including the items that were discarded post-hoc, can be found in the Appendix.

15. Bare *wh*-questions

- a. Non-Crossed Island, Gap/RP

ʔayš	amal	zar-at-/uh	lamman	kan-at
What	amal	visit-3SF-/it	when	was-3SF
fi ar-ryaḍ?				
in the-Riyadh?				

'What did you visit (it) when you were in Riyadh?'

- b. Crossed Non-Island, Gap/RP

ʔayš	tetwaqʕ-een	ʔnn mḥmmad	raḥ	yebeeʕ-/h	fii
what	think-2SF	that Muhammad	will	sell.3SM-/it	in
mḥall-uh al-jdeed?					
shop-his the-new?					

'What do you think that Muhammad will sell in his new shop?'

- c. Crossed Adjunct Island, Gap/RP

ʔayš aš-šurṭah mesik-uu al-ḥarami lamman
what the-police arrested-3PLM the-thief when
əstraq-/=uh min baitu-kum al-jdeed?

stole.3SM-/=it from house-your the-new?

'What did the police arrest the thief when he stole from your new house?'

d. Crossed RC-Island

ʔayš Mḥmmad ʕazam al-fannan illi
what Muhammad invited.3SM the-artist who
rasam-/=uh ʕla jedar al-maṭḥaf?

drew.3SM -/=it on wall the-museum?

'What did Muhammad invited the artist who drew on the museum wall?'

16. Relative clauses

a. Non-Crossed Island, Gap/RP

ḥabai-t al-esbaqatti illi amal sww-at-/=ha
liked-1S the-spaghetti that Amal prepared-3SF-/=it
lamman zur-na-ha fii jeddah
when visited-1PL-her in Jeddah.

'I liked the spaghetti that Amal prepared when we visited her in Jeddah.'

b. Crossed Non-Island, Gap/RP

šef-t as-saḥah illi qulti l-e ʔnn saleḥ
saw-1s the-watch that told-2SF for-me that Saleh
ʔštra-/=ha li-haflat takharruj amal
bought.3SM-/=it for-party graduation Amal

'I saw the watch that you told me that Saleh bought -/it for Amal's graduation party.'

c. Crossed Adjunct Island, Gap/RP

kabbai-t	al-ḥaleeb	illi	khaled meriḍ
spilled-1S	the-milk	that	Khaled felt-sick.3SM
baḥdma	šerib -/=uh	fi al-mṭṣam	ʔms
after	drank.3SM-/=it	in the-restaurant	yesterday

'I spilled the milk that Khaled became sick after he drank in the restaurant yesterday'

d. Crossed RC-Island, Gap/RP

ʔkal-t	al-ḥala	illi	muna	teṣref	al-bent illi
ate-1S	the-dessert	that	Muna	knew.3SF	the-girl that
saww-at-/=uh	fii	al-ḥaflah	ʔms		
made-3SF-/=it	in	the-party	yesterday		

'I eat the dessert that Mona knows the girl who made in the party yesterday'

Importantly, this design makes direct comparison between RTs across different conditions confounded, for the following reasons: (i) experimental sentences are not lexically matched across conditions; (ii) the spillover region in the non-crossed island condition is itself an island, rather than an adjunct phrase. To address our research questions regarding the processing advantage of RPs in bare *wh*-questions, we will depend on differential RTs between RTs in the lexically matched gap dependencies and RP dependencies in each condition, after which we can determine whether these differential RTs are affected as the dependency structures become more complex.

We stated above that relative clauses were included in this experiment to be used as benchmarks for the processing cost induced by islands in the absence of syntactic movement. For this purpose, a direct comparison across RP conditions is required. However, due to the confounds induced by such a comparison, we opted to abandon this methodological choice, which was intended to provide further evidence for the contribution made by processing factors to islands.

It should however be noted here that in addition to the evidence from Study 1, there is also plenty of evidence to support the contribution of processing factors to islands, as we illustrated in Section 3.4 and Section 3.5.1.2.3. Thus, although addressing this issue in the current study would further support the hypothesis that processing factors

contribute to island effects, abandoning this aspect of the experiment does not change the overall argument.

Therefore, in relative clause data, we can only test how the gap-RP manipulation is reflected in RT data as expected by the grammaticality judgement data.

5.2.3 Participants

A total of 50 undergraduate students from Baha University (female section) participated in this experiment; all participants were within the 18-24 age bracket. All were native speakers of Baha Arabic (based on their responses to a language background questionnaire).² Participants received course credit for their participation.

5.2.4 Procedures

In order to satisfy ethical standards of research, full ethical approval was obtained prior to commencing the experiment (reference number: PVAR 17-021). In addition, participants were given and asked to read an information sheet, and provided their informed consent, before taking part in the study.

Participants' working memory capacity was measured, which allowed for the inclusion of this factor as a covariate in the analyses. We did not use the backward digit-span task that we used in the acceptability judgement experiment, as the results of that experiment revealed no association between the linguistic judgements and the WM scores. Hence, we decided to use a different measure for non-verbal WM in the present study. Specifically, the Corsi block-tapping task was used for this purpose (Stoet, 2010; Stoet, 2017). In this task, nine blocks appear on the computer screen. Then, sequences of blocks are marked (lightened), beginning with a sequence of two blocks. Participants were asked to tap the 'marked' blocks in the same sequence they saw. The sequences of marked blocks increased in length as participants proceeded further in the experiment. The largest correct number of marked blocks that the participant remembered was recorded as the participant's score of WM capacity.

Additionally, participants' speed of processing was measured using the digit-symbol coding task (Germine et al., 2012). In this task, a number of symbols, each of which

²This questionnaire is the same as the one used in the acceptability judgement experiment. A translated version of the language background questionnaire can be found in the Appendix.

is associated with a particular number (1 to 3), were presented on the computer screen. Each symbol was then presented in isolation, and participants were asked to press the number associated with this symbol on the keyboard as quickly as possible.

Following the completion of the individual cognitive measures, participants completed the self-paced reading experiment. The experiment was designed using OpenSesame software (Mathôt et al., 2012). Participants were asked to read sentences that were presented word-by-word on the computer screen (one word at a time in the middle of the screen). Each keypress (of the space bar) revealed one word and hid the previous word. Each sentence was followed by a true/false comprehension question to force participants to pay attention while performing the task. See (17).

17. ʔayš tetwaqʕ-een ʔnn Mħmmad raħ yebeeʕ-/h
 what think-2SF that Muhammad will sell.3SM-/it

 spill over fii mħall-uh al-jdeed?
 spill over in shop-his the-new?

‘What do you think that Muhammad will sell spill over in his new shop?’

True/False Question:

‘Muhammad has a new shop.’

There were a total of 132 experimental sentences to read, interspersed with 132 filler sentences, presented across 22 conditions (including the adjunct chains conditions in bare *wh*-questions and the short non-island relative clauses that are discarded in the analysis; i.e. each condition was lexicalised six times). Due to the large number of stimuli involved, there were two sessions, each consisting of three blocks (with 22 experimental sentences and 22 fillers in each block). Each condition was represented with one sentence in each block (six sentences representing the short non-island/island (adjunct chain) *wh*-questions and short non-island relative clauses in each block were abandoned in the statistical analysis). There was one minute break provided in between each block. The design of the blocks ensured that no minimal pairs appeared in the same block. Each session lasted for 30 minutes. The interval between the two sessions for all participants was almost two weeks. Experimental sentences were presented in a different order in each block for each participant; this was done to mitigate any potential effect arising due to repeated exposure to a particular syntactic structure, as well as fatigue effects.

As for the filler sentences, we used a subset of the filler sentences used in the first experimental study. These filler sentences were designed to match with the experimental items in terms of number of clauses and overall length as much as possible. Specifically, all filler sentences are bi-clausal. They included perfectly acceptable and completely unacceptable sentences. Half of the fillers were questions, while the other half were declarative sentences. An example of a grammatical filler sentence is presented in (18). The ungrammatical filler sentences were designed to contain certain grammatical violations, such as subject-verb agreement mistakes (19), sub-categorisation errors (20), and *wh*-questions with incorrect *wh*-words (21).

18. Mona qal-at l-ee ?enn-ha begy-et fii
Mona told-3SF for-me that-she stayed-3SF in
al-mustašfa lhad al-fajr
the-hospital till the-morning
'Mona told me that she stayed in the hospital till the morning.'
19. Laish almdarres-at masha l-beaut-hum badree?
Why the-teacher-PF walked-3SM to-house-their early?
20. Layla qal-at inn Mohammad yeŕtaqed al-kitaab
Layla said-3SF that Mohammad thinks.3SM the-book.
21. Meen akal-tii lamman reḥ-ti l-al-maŕŕam
Who ate-2SF when went-2SF to-the-restaurant
al-jedeed?
the-new?

5.2.5 Results

5.2.5.1 Bare *wh*-questions

In this section, we will conduct two analyses: in the first, we will use differential RTs as our outcome variable, while in the second, we will use non-differential RTs as our outcome variable.

5.2.5.1.1 Differential RT analysis

This first analysis is carried out to address the questions regarding the processing advantage of intrusive RPs in bare *wh*-questions.

- Non-Crossed Island / (RP/GAP)
- Crossed Non-Island / (RP/GAP)
- Crossed RC-Island / (RP/GAP)
- Crossed Adjunct-Island / (RP/GAP)

Since the experimental sentences are not lexically matched across conditions, we calculated the RT differentials between RTs in gap dependencies and RP dependencies in each condition and attempted to determine whether these differential scores are affected as the dependency structures become more complex. The region of interest in the current study is the spillover region, which consists of the three-word adjunct phrases that follow the subcategorisation verb.

We chose this region as our region of interest since comparing RTs in gap dependencies and RP dependencies at the integration site (i.e. the subcategorising verb) will necessarily be longer if there is a pronoun as opposed to no lexical material.

We illustrated above that the spillover region across conditions consists of a three-word adjunct phrase, except in the Non-Crossed Island condition, where the spillover region is an island structure. In terms of processing, this means that parsers are not simply rounding up the meaning of the sentence at the spillover region in this condition (as expected in other conditions), but are also encountering an island clause boundary. However, in the differential RT analysis, we will calculate RT differentials between RP and gap structures in each condition to find out whether the gap vs RP manipulation has had an effect. The impact of encountering an island is controlled in the RP structures and gap structures in the Non-Crossed Island condition, and any difference would be due to the manipulation of the type of element at the dependency tail (RP vs gap). Thus, assuming that any variation in differential RTs in the Non-Crossed Island condition would reflect the impact of the gap vs RP manipulation, as in the other conditions, we decided to retain it in this analysis.

We predict that differential RTs will be affected by the complexity of the host structure (i.e. whether the dependency crosses an island boundary, crosses a non-island clausal boundary or does not cross a clause boundary). Furthermore, we expect the

size of the differential RTs to vary according to participants' WM score and baseline processing speed. We further predict that WM and processing speed scores will be correlated.

Descriptive data

Differential RTs at the spillover region for each minimal pair are calculated as follows: $RT(\text{Gap}) - RT(\text{RP})$. Hence, a positive difference score indicates that gaps are more costly to process, while negative difference score signifies that RPs are more costly in terms of processing. The mean of differential RTs along with 95% Confidence Intervals are plotted in Figure 5-1 below.

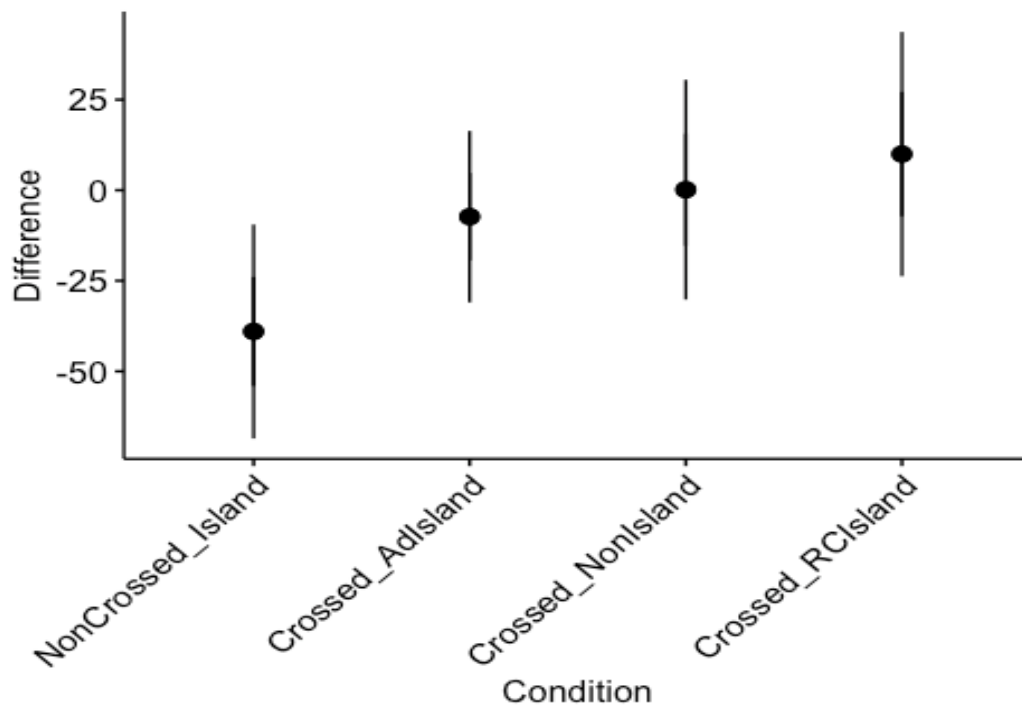


Figure 5-1 Means with confidence intervals of differential RTs at the spillover region across different conditions in bare *wh*-questions.

A positive difference score indicates that gaps are more costly to process, while a negative difference score means that RPs are more costly in terms of processing

The means of differential RTs in bare *wh*-questions in Figure 5-1 suggest that gaps are processed faster than RPs in the Non-Crossed Island condition. In the other conditions (Crossed Adjunct Island, Crossed RC-Island and Crossed Non-Island condition), there seems to be no significant difference between the processing of gap and RP structures. The statistical analysis below will reveal whether or not these results are statistically significant.

Statistical analysis

Data collected in the present study were analysed with Generalised Additive Models (Wood, 2011), using the *mgcv* package (Version 1.8.25) in R Studio (Version 3.5.0). This model was used because it takes into consideration the non-linear nature of reading time data. We further implemented scaled t models to deal with the heavily-tailed data; see Figure 5-2.

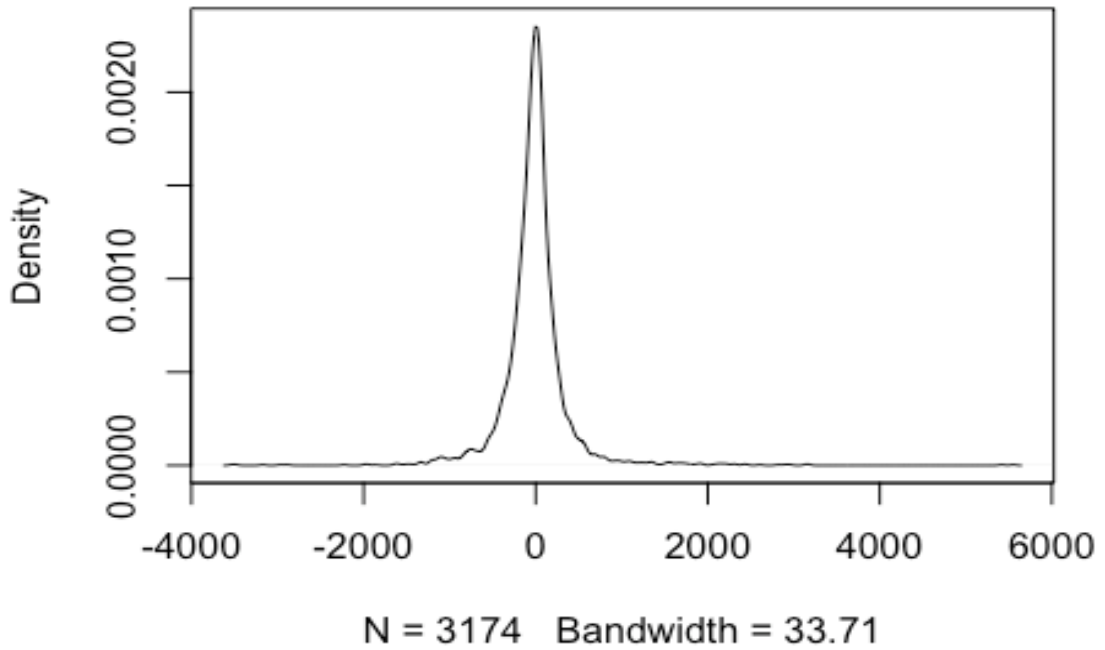


Figure 5-2 Distribution of differential RTs.

The models are fitted starting from random effects only, after which fixed effects are added incrementally. The first model was constructed with Subject, Lexicalisation (which identifies each pair of sentences with or without RP) and Trial (the order in which a sentence is presented to a participant in an experiment) as random factors. The random intercept of Trial was included to mitigate fatigue effects. Subsequently, random slopes and fixed effects were added one by one and retained only if they improved the model fit (as indicated by likelihood ratio comparisons). No participant had been excluded from the statistical analysis based on the mean question-answer accuracy. Accuracy rate was always higher than 80% for both experimental sentences and filler sentences across all subjects. One participant had to be excluded as he did not perform the WM task. To facilitate the interpretation of the model

summary, we scaled and centred the WM variable so that the intercept would align with the mean for that variable.

In the optimal model, the RT differential at the spillover region was predicted by a two-way interaction between Condition (Non-Crossed-Island, Crossed Non-Island, Crossed Adjunct-Island, and Crossed RC-Island) and centered.WM. Random effects included random intercepts for Lexicalisation and random slopes for Subject by Session and Trial by Session (as noted above, the experiment was performed across two sessions by every participant).

Residuals with standard errors < 2.5 were trimmed in the optimal model. A summary of the optimal model is presented in Table 5-1. To facilitate better understanding of the model, coefficients have been plotted in Figure 5-3.

Table 5-1 Summary of the Generalised Additive Model fitted to Difference data. Dependent variable: Difference: RT(gap) – RT (RP).

Coefficients	Estimate	Std.Er	t value	p-value
Intercept	-56.62	36.67	-1.54	0.12
Condition				
NonCrossed_Island	—			
Crossed_NonIsland	38.82	23.61	1.64	0.10
Crossed_AdIsland	24.38	22.46	1.08	0.27
Crossed_RCIsland	29.91	22.46	1.33	0.18
centered.WM				
centered.WM	-2.90	9.30	-0.31	0.75
Condition *centered.WM				

Crossed_NonIsland * centered.WM	15.88	9.43	1.68	0.09
Crossed_AdIsland * centered.WM	2.34	8.96	0.26	0.79
Crossed_RCIsland * centered.WM	8.31	8.98	0.92	0.35

Reference level: Non-Crossed Island. Formula: Difference ~ s(subject, by = session, bs = "re") + s(lexicalisation, bs = "re") + s(trial, by = session) + (Condition * centered.WM)

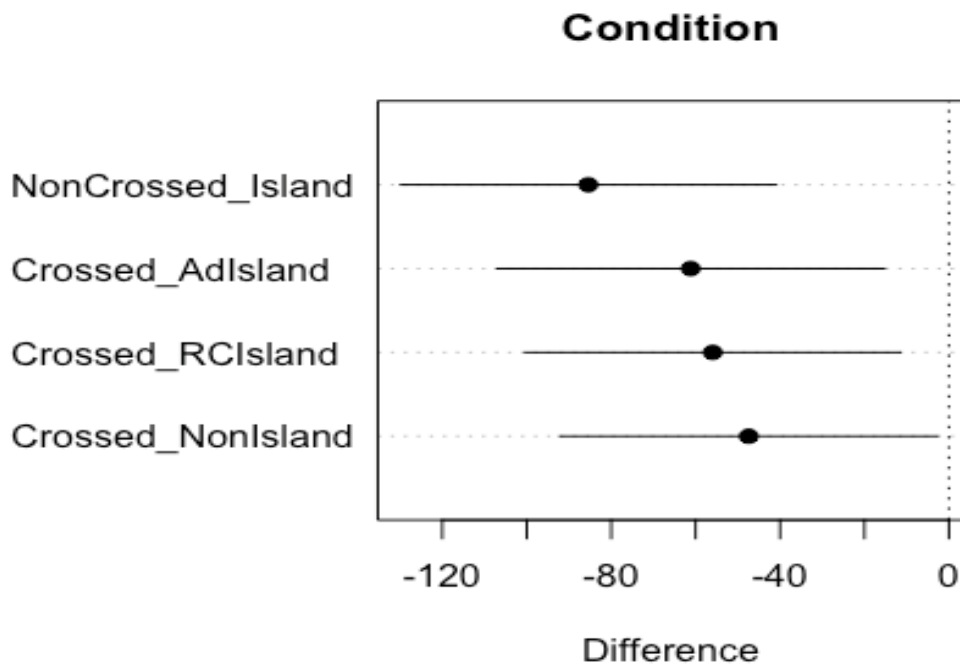


Figure 5-3 The summed effects of the parametric terms in the GAM model.

The summed effects are the predicted response for a certain condition. Accordingly, all the partial effects that apply to these conditions are summed up, including the intercept.

The results revealed no significant difference between the means of differential RTs in the non-crossed island and crossed non-island condition (estimate= 38.82, $t = 1.64$, $p = 0.10$), crossed adjunct-island condition (estimate= 24.38, $t = 1.08$, $p = 0.27$) and crossed RC island condition (estimate= 29.90, $t = 1.33$, $p = 0.18$), suggesting that the processing advantage of intrusive RPs is not sensitive to whether or not a dependency crosses an island/non-island clause boundary.

To determine whether RP structures are processed faster than gap structures in a certain condition, we depend on the estimate of each condition when they occur as the intercept of a model. The intercept estimates revealed no significant difference

between the processing of gap dependencies and RP dependencies in the non-crossed island condition (estimate= -56.62, $t = -1.54$, $p = 0.12$), in the crossed non-island condition (estimate= -29.13, $t = -0.83$, $p = 0.40$), in the crossed Adjunct island condition (estimate= -42.60, $t = -1.24$, $p = 0.21$), and in the Crossed RC island condition (estimate= -37.57, $t = -1.09$, $p = 0.27$).

In summary, our results revealed that no significant difference exists between differential RTs across conditions, although there is a numerical trend suggesting that the differential score is larger in the Non-Crossed Island condition. These results also revealed that there is no significant difference between the processing of structures with gaps and structures with RPs in each condition, although a numerical trend was observed suggesting that gap-structures are processed faster than RP structures across conditions. The findings also indicate that WM scores have no significant impact on results.

5.2.5.1.2 Non-differential RT analysis

Descriptive data

In this analysis, we aim to compare the model with differential RTs as the outcome variable with a model with non-differential RTs as the outcome variable. Our aim is to find out the extent to which results of the non-differential RT-model tally with those of the differential RT model. As discussed above, non-differential RT-analysis might be confounded because the experimental sentences are not lexically matched across conditions. A further confound for such an analysis is that the spillover region in the non-crossed island condition is an island rather than an adjunct phrase.

To carry out a reliable non-differential RT analysis, we decided to include the subset of data for which there is no confounding factor (i.e. we excluded the non-crossed islands). We further performed a check using random effects, as we did in the first acceptability judgement study, to ascertain whether some items behave differently. The means of raw RTs with 95% Confidence Intervals of RTs in gap bare wh-questions and RP bare wh-questions are plotted in Figure 5-4 below.

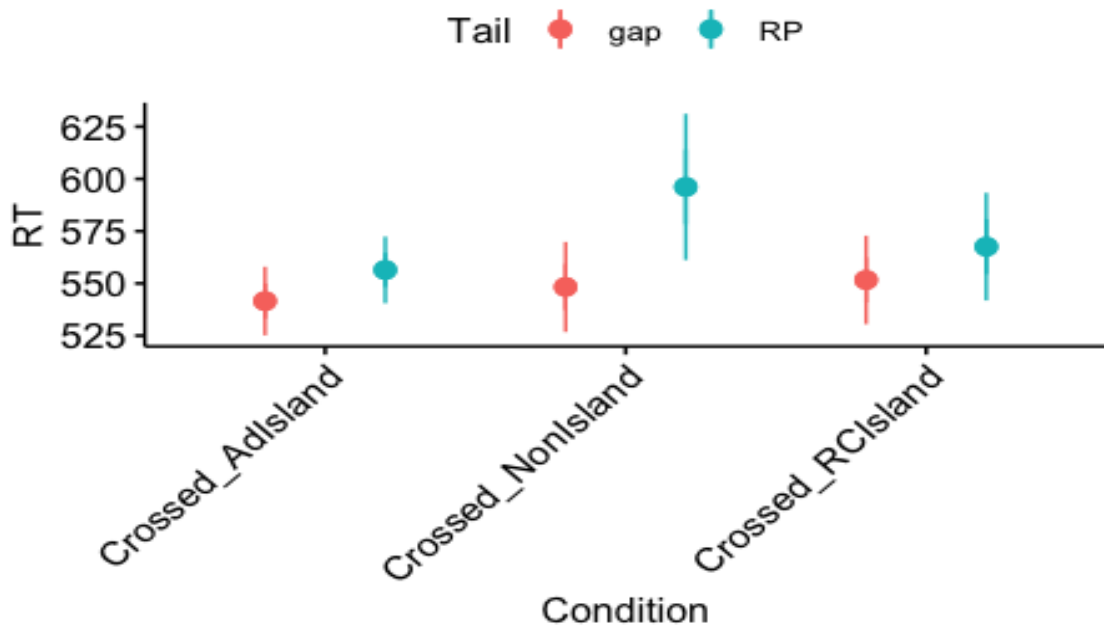


Figure 5-4 Means with confidence intervals of raw RT scores in gap-bare *wh*-questions and RP-bare *wh*-questions across conditions.

The means of raw RT scores in Figure 5-4 suggest that structures with RPs tend to be processed more slowly than structures with gaps. The statistical analysis below will reveal whether or not these results are statistically significant.

We stated above that in this study we will not focus on the subcategorizing verb as our region of interest since comparing RTs in gap dependencies and RP dependencies at this region will necessarily be longer if there is a pronoun as opposed to no lexical materials. Instead, we evaluate reading times at the spillover region following the RP or gap, which comprises an adjunct phrase consisting of three words. However, for clarity sake, Table 5-2 shows the mean raw RTs at integration site and the mean average raw RTs at spillover region across conditions.

Table 5-2 The mean raw RTs at integration site and the mean average raw RTs at spillover region (with standard deviations) across conditions.

Variable 1	Variable 2	RT at the verb region		Average RT at spill over region	
		RT.mean	RT.sd	RT.mean	RT.sd
Tail	Condition				

Gap	Crossed Adjunct Island		541.08	345.65	539.76	278.07
Gap	Crossed Non- Island		569.59	330.85	543.22	300.76
Gap	Crossed RC Island		519.77	268.37	555.49	374.97
RP	Crossed Adjunct Island		598.65	431.52	541.39	239.32
RP	Crossed Non- Island		702.46	509.02	539.18	294.28
RP	Crossed RC Island		648.43	570.89	545.24	318.87

Statistical analysis

Prior to statistical analysis, raw RT scores with a value greater than 5000 ms and smaller than 150 ms were removed from the data. These values were chosen because values under 150 ms likely indicate that participants did not read the word (whether accidentally or on purpose), while values over 5000 ms indicate that parsers were distracted by a non-controlled factor. Eleven data points were removed as a consequence of this procedure: six points from the gap conditions, and five points from the RP conditions.

Prior to statistical analysis, raw reading times were box-cox transformed to normalise the data. The distribution of box-cox RT data is plotted in Figure 5-5.

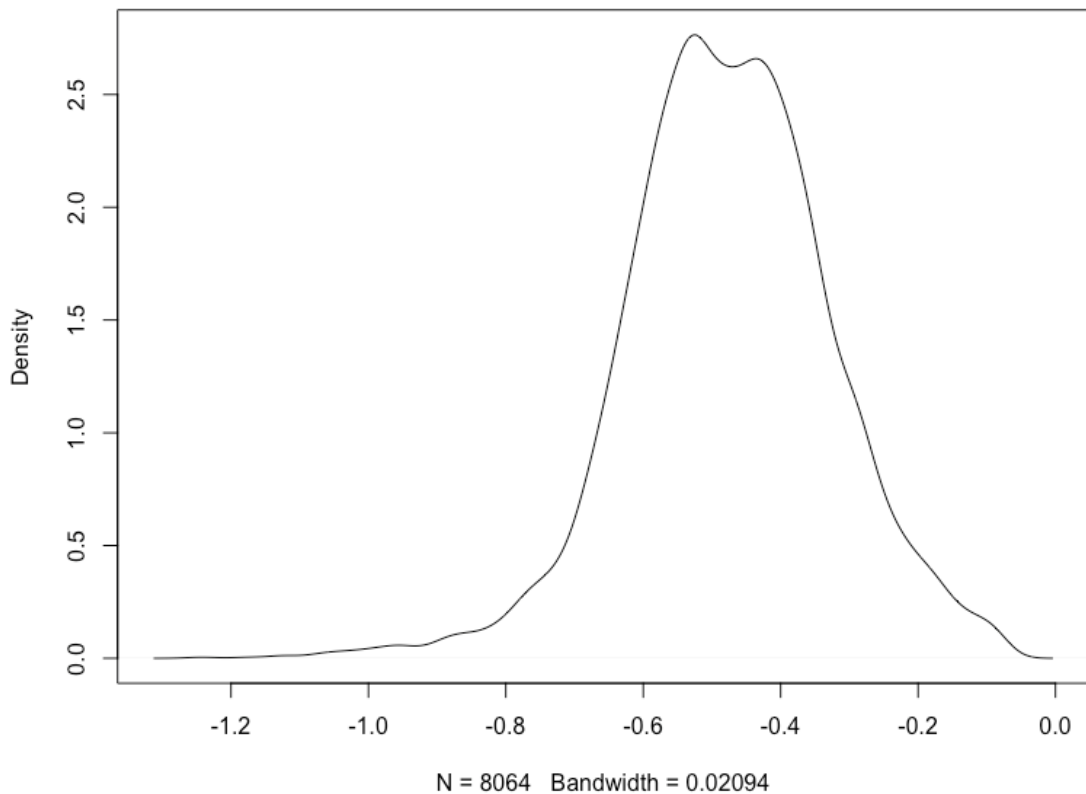


Figure 5-5 Distribution of box-cox non-differential RT data.

As the plot in Figure 5-5 suggests, these data are not normally distributed. Given the non-normal distribution, we opted to fit a non-parametric additive quantile regression model to the RT data (Fasiolo et al., 2017), using the *mgcViz* package (Version 0.1.1) in R Studio (Version 3.5.0).

The models are fitted starting with random effects only, after which fixed effects are added incrementally. The first model was constructed with Subject, Trial (the order in which a sentence is presented to a participant in an experiment) and Lexicalisation (which identifies each pair of sentences with or without RP) as random factors. Subsequently, random slopes and fixed effects were added one by one and retained only if they improved the model fit, as indicated by likelihood ratio comparisons.

In the optimal model, RT scores at the spillover region were predicted by Tail (gap vs RPs). Adding the Condition variable did not improve model fit, either as a main effect or interaction. Similarly, adding the centered.WM variable to the model did not improve model fit, either as a main effect or interaction. Random effects included random intercepts for Lexicalisation and Trial and random slope for Subject by Session.

The model was fitted on the following quantiles: 0.1, 0.25, 0.5, 0.7 and 0.9. Results of the models on different quantiles revealed no significant impact of Tail on processing, except for quantile '0.25'. The results of this model revealed that structures with gaps are processed faster than structures with RPs (estimate= 0.009, t= 3.083, p= 0.002). This means that the disruption of processing RPs in bare *wh*-questions is only observable in the shorter reading times (i.e. not for slow readers). A summary of this model is presented in Table 5-3. Data are plotted in Figure 5-6.

Table 5-3 Summary of the non-parametric additive quantile regression model of RT data: quantile '0.25'.

Dependent variable: box-cox RT. Reference level: gap. Formula: RT.t ~ s(subject, by = session, bs = "re") + s(lexicalisation, bs = "re") + s(trial) + (Tail)

Coefficients	Estimate	Std.Err	t value	p-value
Intercept	-0.55	0.01	-55.147	0.0001
Tail				
RP	0.009	0.003	3.089	0.002

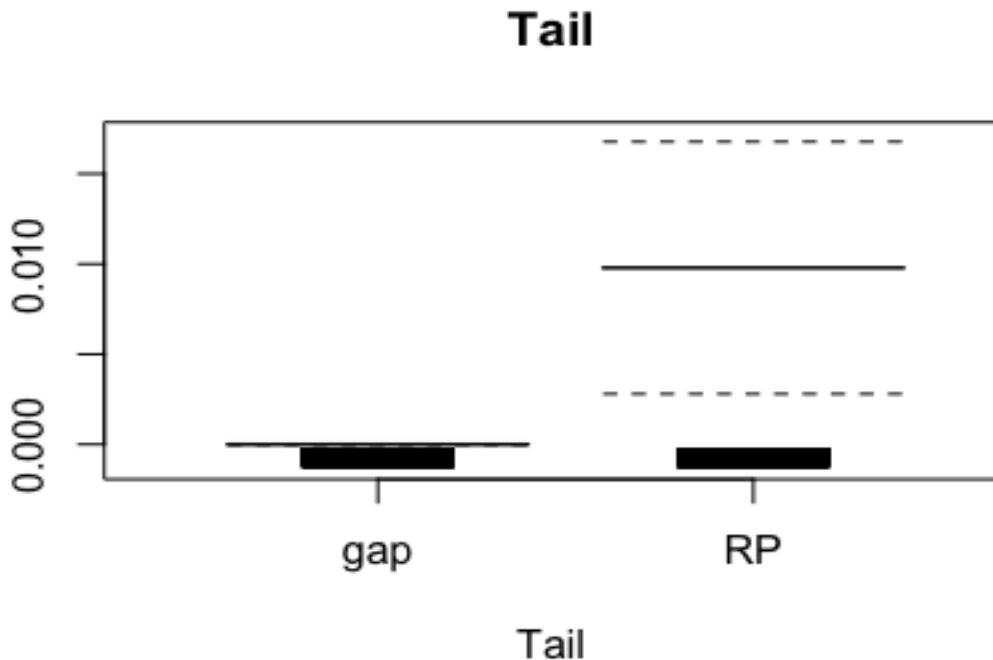


Figure 5-6 Predicted values of RT scores for Tail in bare *wh*-questions. Reference level: gap.

Since the experimental sentences across the variable Tail are lexically matched, we have no concerns regarding the robustness of the statistical analysis. That is, there is no need to perform a check (using random effects) as we did in the first acceptability judgement study in order to ascertain whether some items will behave differently.

We re-ran the analysis with the Non-Crossed Island condition included. Similar to the previous analysis, the optimal model was found to be the one where the outcome variable was predicted by Tail. Adding Condition and centered.WM to the model did not improve model fit either as main effects or interactions. The results of this model revealed that gap structures are processed significantly faster than RP structures in the following quantiles: quantile 0.1 ($p=0.001$), quantile 0.25 ($p= 0.000$), quantile 0.5 ($p= 0.000$) and quantile 0.7 ($p=0.002$). In quantile 0.9, there was no significant difference between the processing of gap and RP structures ($p= 0.11$). This means that the disruption of processing RPs in bare *wh*-questions is not observable in longer reading times (i.e. for slow readers).

5.2.5.2 Relative clauses

The aim of this analysis is to investigate the extent to which RP structures in Baha Arabic relative clauses are processed faster than gap structures, as predicted by the pattern of their acceptability in the grammaticality judgement task. Since we are not concerned with investigating this effect across conditions (as acceptability judgement revealed that resumption in *illi*-structures is obligatory regardless of condition), we can use non-differential RT as the outcome variable in this analysis, and the Non-Crossed Island condition can be included.

5.2.5.2.1 Descriptive data

The mean of raw RTs with 95% Confidence Intervals of RTs in gap relatives and RP relatives are plotted in Figure 5-7.

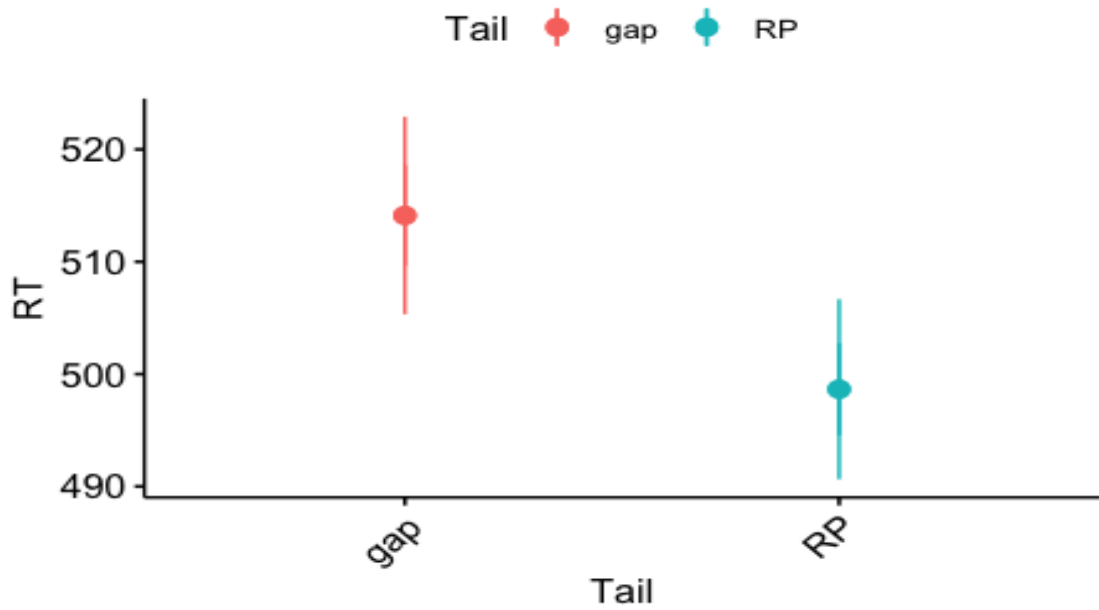


Figure 5-7 Means with confidence intervals of raw RT scores in gap relatives and RP relatives.

The means of raw RT scores in Figure 5-7 suggest that relatives with RPs are processed faster than relatives with gaps. The statistical analysis below will reveal whether or not these results are statistically significant.

We stated above that in this study we will not focus on the subcategorizing verb as our region of interest since comparing RTs in gap dependencies and RP dependencies at this region will necessarily be longer if there is a pronoun as opposed to no lexical materials. Instead, we evaluate reading times at the spillover region following the RP or gap, which comprises an adjunct phrase consisting of three words. However, for clarity sake, Table 5-4 shows the mean raw RTs at integration site and the mean average raw RTs at spillover region across conditions.

Table 5-4 The mean raw RTs at integration site and the mean average raw RTs at spillover region across conditions.

Variable	RT at the verb region		Average RT at spill over region	
	RT.mean	RT.sd	RT.mean	RT.sd
Gap	498.49	248.67	515.77	272.73

RP	544.26	352.94	501.94	293.73
----	--------	--------	--------	--------

5.2.5.2.2 Statistical analysis

Prior to statistical analysis, raw RT scores with a value greater than 5000 ms and smaller than 150 ms were removed from the data. These values were chosen because values less than 150 ms likely indicate that participants did not read the word (either accidentally or on purpose), while values over 5000ms indicate that parsers were distracted by a non-controlled factor. Four data points were removed as a consequence of this procedure: three points from the gap conditions and one from the RP conditions.

Prior to statistical analysis, raw reading times were box-cox transformed to normalise the data. The distribution of box-cox RT data at the spillover region is plotted in Figure 5-8.

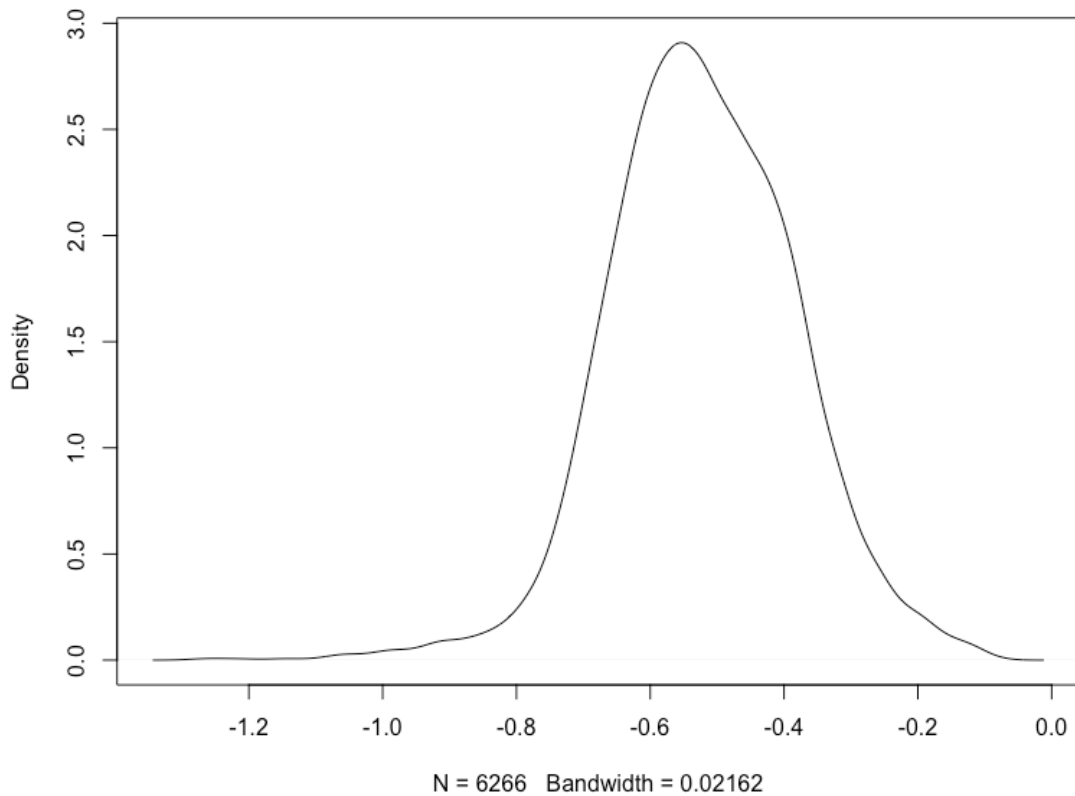


Figure 5-8 Distribution of box-cox transformed RTs at the spillover region

Modelling procedures

Data collected in the present study were analysed with Generalised Additive Models (Wood, 2011), using the mgcv package (Version 1.8.25) in R Studio (Version 3.5.0). This model was used because it takes into consideration the non-linear nature of reading time data.

The models are fitted starting with only random effects, after which then fixed effects are added incrementally. The first model was constructed with Subject, Trial (the order in which a sentence is presented to a participant in an experiment), Lexicalisation (which identifies each pair of sentences with or without RP) and WordsLength (the number of letters in the spillover region) as random factors. Afterwards, random slopes and fixed effects were added one by one and retained only if they improved the model fit, as indicated by likelihood ratio comparisons.

In the optimal model, the non-differential RT score at the spillover region was predicted by Tail (gap vs RPs). Random effects included random intercepts for Lexicalisation, Trial and WordsLength and random slope for Subject by Session. A summary of the optimal model is presented in Table 5-5.

Table 5-5 Summary of the Generalised Additive Model fitted to Difference data.

Dependent variable: box-cox RT. Reference level: gap. Formula: $RT.t \sim s(\text{Subject}, \text{by} = \text{Session}, \text{bs} = \text{"re"}) + s(\text{Lexicalisation}, \text{bs} = \text{"re"}) + s(\text{Trial}, \text{bs} = \text{"re"}) + s(\text{WordsLength}, \text{bs} = \text{"re"}) + (\text{Tail})$

Coefficients	Estimate	Std.Err	t value	p-value
Intercept	-0.441	0.011	-37.960	0.0001
Tail				
RP	-0.006	0.002	-2.663	0.006

Results revealed a significant difference between the processing of relatives with gaps and those with RPs, such that relatives with RPs are processed significantly faster than gaps at the spillover region (estimate= -0.006, t= -2.66, p= 0.006).

5.2.6 Discussion

The current experimental study investigates the extent to which intrusive RPs in bare *wh*-questions facilitate the processing of complex structures in Baha Arabic. Based on the results of the acceptability judgement experiment, we contended that bare *wh*-questions are movement dependencies with intrusive RPs. Hence, we predicted that intrusive RPs in bare *wh*-questions will interact with whether the dependency structure crosses a non-island/island clause boundary or not. We further predict that the facilitative effect of intrusive RPs will be more evident in dependencies crossing weaker island boundaries (i.e. adjunct islands) than in dependencies crossing stronger island boundaries (i.e. RC islands).

As for relative clauses, the results of the acceptability judgement experiment revealed that these are binding dependencies with obligatory true RPs; hence, no interaction with the processing complexity of structures is expected. However, we included these data to examine the RT correlates of the grammaticality judgements elicited in Study 1.

In brief, the results of the current study yield no evidence that RPs have a processing effect in bare *wh*-questions. More specifically, bare *wh*-questions with RPs are associated with longer reaction times in comparison to bare *wh*-questions with gaps across conditions (particularly for faster readers, as suggested by the outcomes of the non-parametric additive quantile regression model). The longer RTs observed in bare *wh*-questions with RPs suggest that RPs are ungrammatical in bare *wh*-questions. Intrusive RPs seem to be unexpected across conditions, but tend to be better tolerated (i.e. create less of a surprise effect) as processing complexity increases, as suggested by the numerical trend in the differential RT analysis. We assume that intrusive RPs in bare *wh*-questions might have a processing function in complex structures, but some aspects of the experimental design might not allow this facilitation effect to be detected. These methodological issues will be highlighted in the limitations section below.

Moving on to the impact of RPs on the processing of relative clauses, our results revealed that, as expected by the grammatical account, structures with RPs are processed faster than structures with gaps. This result provides further evidence to support the argument that the function of RPs in Baha Arabic relatives is different from that of RPs in Hebrew (viz., Friedmann and Costa, 2011). Meltzer-Asscher (2018) concluded that, after reviewing the results of several online experimental studies on

Hebrew RPs, RPs in Hebrew hinder processing rather than facilitate it, leading her to propose that resumption in Hebrew is a processing phenomenon that helps the producer, not the comprehender. By contrast, in Baha Arabic, RP structures are processed faster than gap structures, as the reader anticipates their presence (on account of syntactic licensing). This suggests that that RPs in relative clauses are intrusive in Hebrew but required for grammatical reasons in Baha Arabic.

5.2.7 Limitations

The acceptability experiment on Baha Arabic (in Study 1) revealed that RPs in Baha Arabic are marginally accepted in non-island contexts in bare *wh*-questions. However, the acceptability pattern was not consistent. Variability within participants was found such that almost all participants used the full range of acceptability rating for RP dependencies in non-crossed island conditions (i.e. highly acceptable, highly unacceptable, or somewhere in between). We hypothesised that RPs are ungrammatical in *what*-questions in Baha Arabic, but marginally accepted due to the availability of RPs in RCs. We further speculated that intrusive RPs in *what*-questions might facilitate processing, but this could not be determined based on the offline acceptability judgement data.

Hence, we decided to use online methods to test the processing effect of RPs in bare *wh*-questions. However, as explained above, it emerged from the online study that RPs in bare *wh*-questions were never processed faster than gaps across conditions (as revealed by the differential and non-differential RT analyses). These results are not consistent with the hypothesis that RPs facilitate processing.

However, there are some limitations of the present study that might hinder the facilitation effect of RPs in bare *wh*-questions. Importantly, experimental sentences were presented to participants without preceding context. As noted by Hofmeister (2012), presenting direct questions without preceding contexts might increase the processing cost of such structures by making the questions pragmatically odd. This, in turn, might further hinder the emergence of any facilitation effect of intrusive RPs. In other words, in the absence of context, the facilitation effect of RPs might not be detectable.

It is also possible that RPs are simply never allowed in *what*-questions, as is commonly assumed to be the case for many varieties of Arabic, such as MSA and Lebanese Arabic (Aoun et al., 2009). Rather, RPs are allowed in questions containing

'who' and discourse-linked fillers (i.e. *which-N* fillers) in both languages. For instance, the examples below show that RP dependencies are not allowed in MSA with *maḡḡaa* ('what') (22a), but are allowed with *ʔayya kitaabin* ('which book') (22b) and *man* ('who') (22c). Similarly, RP dependencies are not allowed in Lebanese Arabic with *šu* ('what') (23a), but are acceptable with *ʔayya mmasil* ('which actor') (23b) and *miin* ('who') (23b).

22. MSA

a. maḡḡaa_i ʔištara_t(*-hu_i) laila min al-maktabati?
 what_i bought.3fs(*-it_i) Laila from the-bookstore
 'What_i did Laila buy (*it_i) from the bookstore?'

b. ʔayya kitaabin_i ʔištara_t(-hu_i) laila min
 which book_i bought.3fs(-it_i) Laila from
 al-maktabati?
 the- bookstore
 'Which book_i did Laila buy (it_i) from the bookstore?'

(Tucker et al., 2019)

c. man_i zaara_t-hu_i naadia?
 Who_i visited.3fs-him_i Nadia?
 Who did Nadia visit?

(Aoun et al., 2009, p.132)

23. Lebanese Arabic

a. šu_i štaray_t-i_i b-l-maktabe?
 what_i bought.2SF-it_i from the-bookstore?
 'What did you buy from the bookstore?'

(Abdel Razaq, 2011)

b. ʔayya mmasil_i šeft-o_i b-l-maṭṭam?
 which actor_i saw.2MS-him_i in-the-restaurant
 'Which actor did you see in the restaurant?'

(Aoun et al., 2009, p.128)

c. miin_i šeft-o_i b-l- maṭṭam?

who_i saw.2SM-him_i in-the-restaurant

'Who did you see in the restaurant?'

(Abdel Razaq, 2011)

Moreover, Tucker et al. (2019) tested the ameliorative effect of RPs in *wh*-questions with 'what' and *wh*-questions with D-linked fillers (such as 'which book') in MSA. Resumption was strongly penalised in 'what' questions, regardless of whether or not the dependency crossed an island boundary. These authors explained this with reference to the fact that *wh*-fillers such as 'what' are not easily linked to resumptive pronouns. On the assumption that RPs are genuine pronouns (Erteschik-Shir, 1992; Frazier and Clifton, 2002; Aoun et al., 2009; Alexopoulou and Keller, 2013), they are inherently D-linked, and can only be linked with D-linked fillers. Tucker et al. (2019) therefore investigated the impact of RPs in *wh*-questions with inherently D-linked fillers (i.e. *which*-N) and observed an ameliorative effect; specifically, utilising the superadditivity paradigm, Tucker et al. (2019) observed that the superadditive effect of islands is reduced when D-linked *wh*-dependencies terminate in RPs compared to when they terminate in gaps, an effect that is not observed in non-D-linked *wh*-questions. When interpreted in that light, the lack of an ameliorative effect of RPs in the present study might be a result of the non-D-linked nature of the fillers.

Hence, we decided to re-test the processing effect of RPs in bare *wh*-questions in another self-paced reading task, with the following two modifications: (i) experimental sentences are contextualised (to increase the likelihood of the emergence of the facilitative effect of RPs if they do have a facilitation effect in what-questions); (ii) inherently discourse-linked questions are compared with non-inherently-D-linked questions.

Relative clauses are excluded from the next experimental study, as the results of the first and second experiments support the hypothesis that these RPs are required for grammatical purposes.

5.3 Study 3: Processing contextualised D-linked/non-D-linked *wh*-questions with RPs

5.3.1 Rationale

We used a self-paced reading task to retest the processing effect of RPs in *wh*-questions in Baha Arabic. We used the same experimental sentences as in the previous experiment on bare *wh*-questions, but this time, with context sentences preceding each of them. These context sentences are used for two purposes. First, the context enhances the ecological validity of the design by increasing the plausibility of the *wh*-questions, thereby making their processing more natural. Second, the context allows us to control for the discourse status of the referent of the filler phrases (and associated RPs) by ensuring they are not mentioned before the *wh*-question. The contexts included no mention of the referent in question, but instead introduced an element from which this referent could be inferred (i.e. via bridging or association). The context sentences were also designed to maximise the relevance of the following question. This approach allows us to investigate the extent to which intrusive RPs have a processing effect regardless of whether or not the filler phrase is interpreted as discourse-linked.

Furthermore, we included inherently D-linked *wh*-questions (i.e. *which-N* questions) in the current experimental study, as the literature suggests that resumption is more common in this type of question than in other types. WM and processing speed measures were also included to investigate the extent to which individual cognitive differences affect the processing of RP dependencies in these two types of questions.

5.3.2 Materials

Similar to the previous experimental studies, the current study manipulates the type of element appearing at the dependency tail (gap vs RP), the type of dependency (non-crossed island, crossed non-islands, crossed adjunct island, crossed RC island), and the d-linking status of the filler in *wh*-questions – i.e. a 2 x 4 x 2 design. A sample of experimental sentences is presented below. The parts in bold represent the spillover region (i.e. our region of interest).

24. Bare (non-D-linked) *wh*-questions

a. Non-Crossed Island, Gap/RP

Context:

'The pupils in my literature class are all very motivated this year.'

Experimental sentence:

ʔayš raḥ amal tulq-i-/=h law šarak-at
what will Amal recite.3SF-/=it if participated-3SF
fii al-musabaq al-ʔdabyyah?
in the-competition the-literature?

'What would Amal recite if she participated in the literature competition?'

b. Crossed Non-Island, Gap/RP

Context:

'I don't know when I can visit Muhammad to congratulate him for his new business.'

Experimental sentence:

ʔayš tetwaqʕ-een ʔnn mḥmmad raḥ yebeeʕ-/=h fii
what think-2SF that Muhammad will sell.3SM-/=it in
mḥall-uh al-jdeed?
shop-his the-new?

'What do you think that Muhammad will sell in his new shop?'

c. Crossed Adjunct Island, Gap/RP

Context:

'Praise be to Allah that most of the rooms in the house were closed when the thief was inside the house.'

Experimental Sentence:

ʔayš aš-šurṭah mesik-uu al-ḥarami lamman
what the-police arrested-3PLM the-thief when
əstraq-/=uh min baitu-kum al-jdeed?
stole.3SM-/=it from house-your the-new?

'What did the police arrest the thief when he stole from your new house?'

d. Crossed RC-Island

Context:

'This company has a very strange staffing policy, that doesn't take efficiency into account.'

Experimental Sentence:

ʔayš al-mudeer šakar al-muhandis illi
what the-manager thanked.3SM the-engineer who
šammam-/uh fii muddah qašeerah?
designed.3SM-/it in **period short?**

'What did the manager thank the engineer who designed in short period?'

25. D-linked questions

a. Non-Crossed Island, Gap/RP

Context:

'Most of my sisters love Asian food.'

Experimental sentence:

ʔayy maṭṭam šini ʔkhwat-ek za-raw-/h
which restaurant Chinese sisters-your visted-3PL-/it
lamman **kanaw fii jeddah?**
when were in Jeddah?

'Which Chinese restaurant do your sisters visit when they go to Jeddah?'

b. Crossed Non-Island, Gap/RP

Context:

'My brother is still working in the engineering office in Makkah.'

Experimental sentence:

ʔayy vella qul-ti l-e inn akhu-k
which palace told-2SF for-me that brother-your
šamma-/=ha **li-mudeer šarekat al-esment?**
designed.3SM-/=it for-manager firm the-cement?
'Which palace did you tell me that your brother designed for the
cement firm's manager?'

c. Crossed Adjunct Island, Gap/RP

Context:

'The social worker came to visit Ahmad in prison yesterday.'

Experimental sentence:

ʔayy šuruṭ raḥ yefrej-uun ʕan aḥmad
which conditions will release-3PLM about Ahmad
eḍa ejtaz-/=ha **fii settat ʔšhur?**
if passed.3SM-/=it in six months?

'Which conditions will they release Ahmad if he met after six months?'

d. Crossed RC-Island, Gap/RP

Context:

'Policemen are everywhere to reduce disorder.'

Experimental sentence:

ʔayy sayyarah al-ḥarami qatal al-walad
which car the-thief killed-3SM the-boy
illi kan ysuq-/=ha **fii wasaṭ**
who was.3SM driving.3SM-/=it **in centre**
al-madeenah?
the-city?

'Which car did the thief kill the man who was driving in the city centre?'

5.3.3 Participants

A total of 50 undergraduate students from Baha University participated in this experiment; all participants were within the 18-24 age bracket. All were native speakers of Baha Arabic (based on their responses to a language background questionnaire). Participants received course credit for their participation.

5.3.4 Procedures

In order to satisfy ethical standards of research, full ethical approval was obtained prior to commencing the experiment (reference number: PVAR 17-074). The same procedures are employed here as for the previous experimental study. In summary, participants' working memory capacity was measured using the Corsi block-tapping task, after which participants' speed of processing was measured using the digit-symbol coding task. Following the completion of these individual cognitive measures, participants were provided with instructions and practice trials prior to performing the self-paced reading experiment.

The experiment was designed using OpenSesame software (Mathôt et al., 2012). Participants were asked to read sentences that were presented word-by-word on the computer screen (one word at a time in the middle of the screen). Each keypress (of the space bar) revealed one word and hid the previous word. Each sentence was followed by a true/false comprehension question to ensure that participants had truly attempted to understand the dependency structure. An example of an experimental item with a follow-up comprehension question is provided in (26).

26. Context:

'The faculty staff complains about many problems in the department.'

Experimental sentence:

ʔayš tetwaqaʕeen inn al-mudeerah raḥ
what think.2F that the-administrator will
tenaqiš-/uh fi ejtemaʕ al-edarah?
discuss-/it in meeting the-administration?

'What do you think that the administrator will discuss in the administration meeting?'

Comprehension question:

'The faculty staff had no complaints to discuss in the administration meeting'

There were a total of 96 (48 non-discourse-linked *wh*-questions and 48 discourse-linked questions) experimental sentences to read, presented across 16 (eight for each structure) conditions (i.e. each condition is lexicalised six times), and interspersed with 96 distractor sentences (1:1 ratio). Each sentence was preceded by a context sentence and followed by a true/false comprehension question. Experimental sentences were presented in a different order for each participant; this was done to mitigate any potential facilitation effect that might have had an impact on results due to repeated exposure to particular syntactic structures, as well as to alleviate potential fatigue effects. Due to the large number of stimuli and filler sentences involved, there were two experiments, each consisting of three blocks (16 experimental sentences in each block). A one-minute break was provided in between blocks. The design of the blocks ensured that no minimal pairs appeared in the same block. Each session lasted for about 30 minutes.

5.3.5 Analysis and predictions

Assuming that intrusive RPs have a processing function, we predict that differential RTs, calculated as $RT(\text{Gap}) - RT(\text{RP})$, will be more positive as the dependency structure becomes more complex (i.e. in crossed non-island, crossed adjunct-island and crossed RC island conditions), both in D-linked and non-D-linked *wh*-questions. Furthermore, we predict that participants with smaller WM scores and/or processing speed scores will benefit more from RPs in complex conditions (i.e. in crossed non-island, crossed adjunct island and crossed RC island conditions) in comparison to participants with higher WM and/or processing speed scores.

5.3.5.1 Descriptive data

Differential RTs, calculated as the difference in RT at the spillover region in each gap/RP minimal pair across participants, were used as the dependent variable. The spillover regions consist of a three-word adjunct phrase that follows the subcategorisation verb (except in the non-crossed island condition, where the three words in the spillover region include an island clause boundary). Prior to statistical analysis, differential RTs greater than 2000 ms or less than -2000 ms were removed

from the data; these values were removed on basis of the observed visual distribution of the data. See Figure 5-9.

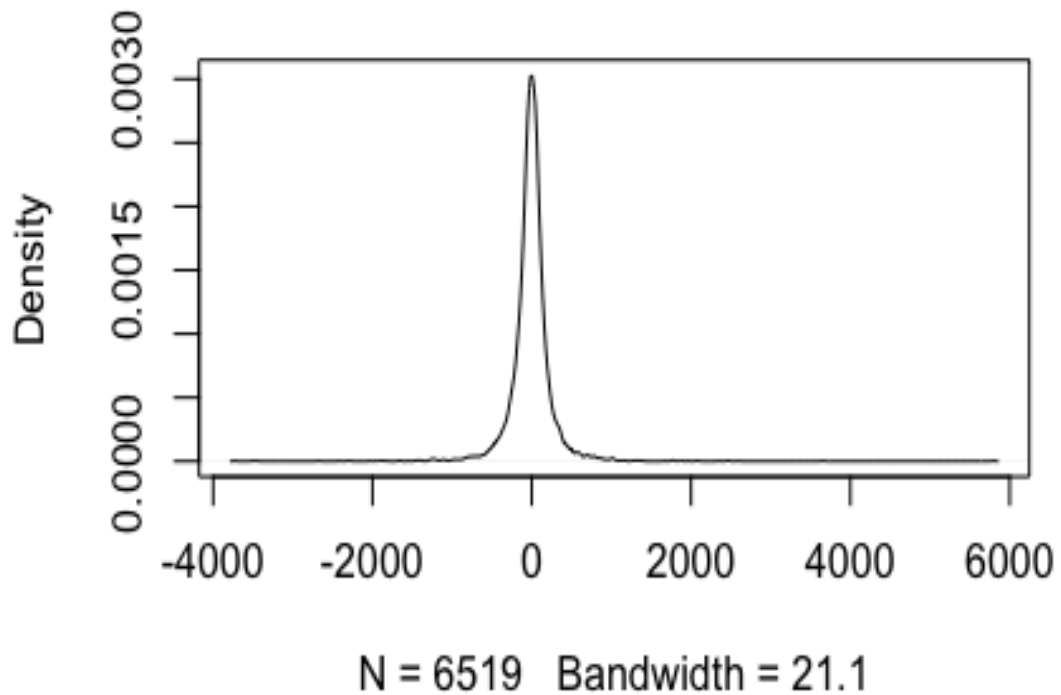


Figure 5-9 Distribution of differential RTs before removing points with an absolute value greater than 2000 ms.

Sixteen data points were removed as a consequence of this procedure: five points from the non-crossed island condition, one point from the crossed non-island condition, four points from the crossed adjunct-island condition and six points from the crossed RC-island condition. The adjusted distribution of differential RT data is shown in Figure 5-10.

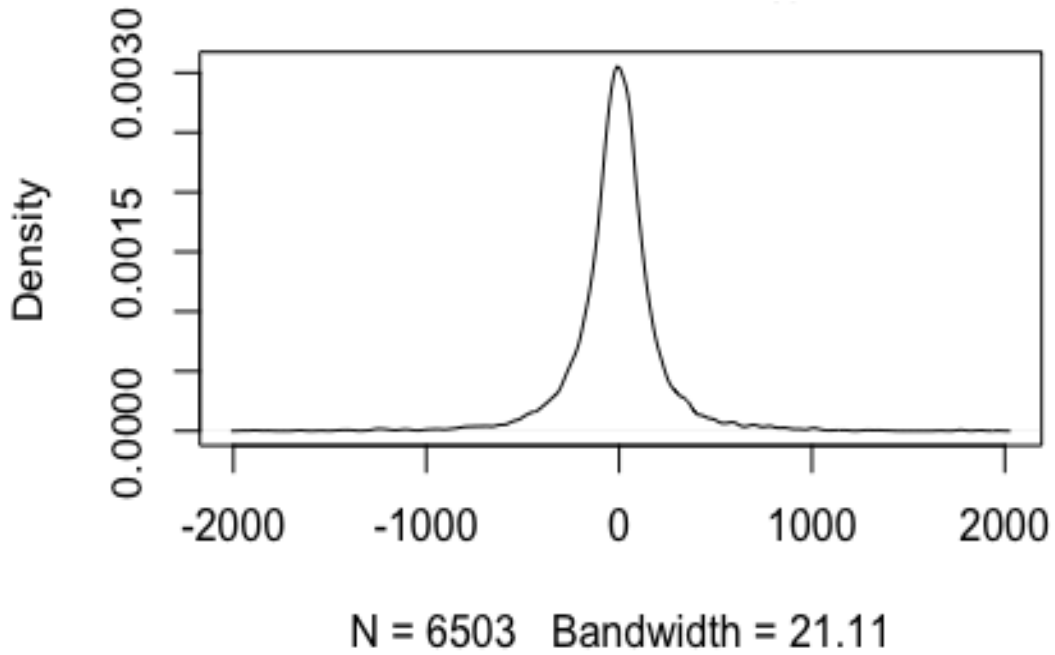


Figure 5-10 Distribution of differential RTs after removing points with an absolute value greater than 2000 ms.

Differential RTs for each minimal pair are calculated as follows: RT (Gap) – RT (RP). Hence, positive RT differentials indicate that gaps are more costly to process, while negative differential RTs means that RPs are more costly in terms of processing. The means of differential RTs and their standard deviations are presented in Figure 5-11.

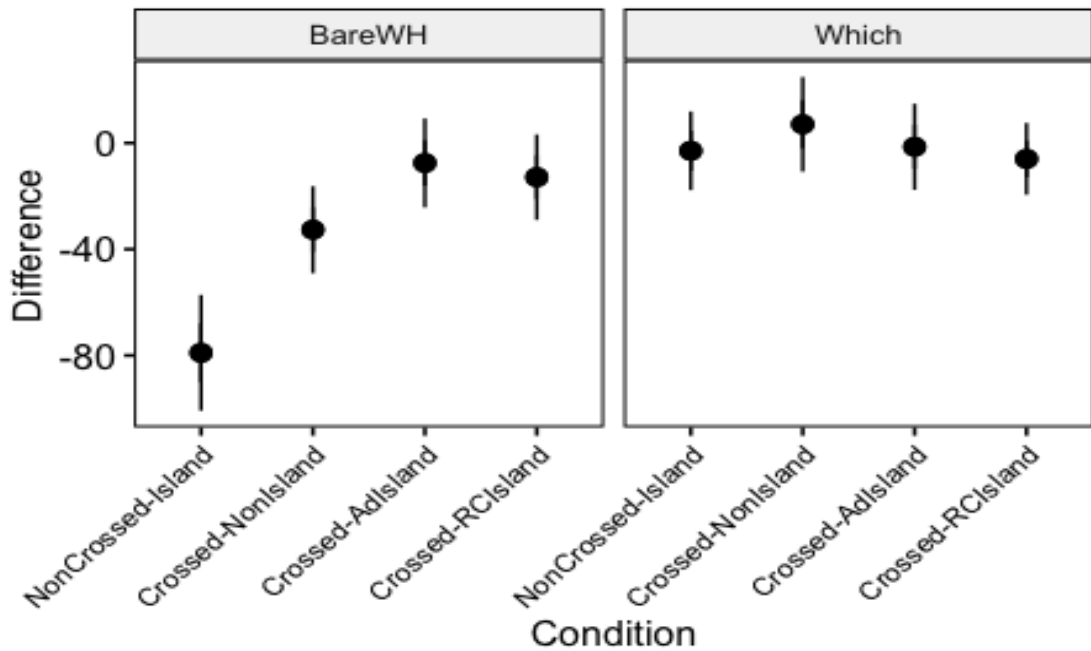


Figure 5-11: Means with confidence intervals of differential RTs across different conditions in D-linked and non-D-linked questions.

Positive differential RTs indicate that gaps are more costly to process, while negative differential RTs means that RPs are more costly in terms of processing.

The means of differential RTs in Figure 5-11 suggest that, in non-D-linked *wh*-questions, structures with gaps are processed faster than structures with RPs across conditions, and moreover that the disruption effect of RPs is more evident in the Non-Crossed Island condition. As for D-linked *wh*-questions, the means of differential RTs suggest that gaps and RPs are processed at a similar rate across conditions. The statistical analysis below will reveal whether or not these results are statistically significant.

5.3.5.2 Statistical analysis

Data collected in the current study were analysed with Generalised Additive Models (Wood, 2011), using the *mgcv* package (Version 1.8.25). in R Studio (Version 3.5.0). This model was used because it takes into consideration the non-linear nature of reading time data. We implemented scaled t models to deal with heavily tailed data.

The models were fitted starting from random effects only, after which fixed effects were added incrementally. The first model was constructed with Subject, Lexicalisation (while identifies each pair of sentences with or without RP), and Trial (the order in which a sentence is presented to a participant in an experiment) as random effects. The random intercept of Trial was included to capture any fatigue or training effect. Subsequently, random slopes and fixed effects were added one by one and retained only if they improved the model fit, as indicated by likelihood ratio comparisons.

Prior to the statistical analysis, one participant had to be excluded because he did not perform the WM task. Furthermore, we scaled and centred the WM variable, since this numeric variable has no value that equals zero; this was done to increase the ease of interpreting the model. No participant had been excluded from the statistical analysis based on the mean question-answer accuracy. Accuracy rate was always higher than 80% for both experimental sentences and filler sentences across all subjects.

In the optimal model, the differential RTs at the spillover region were predicted by a two-way interaction between Structure (non-D-linked vs D-linked *wh*-questions) and Condition (Non-Crossed-Island, Crossed Non-Island, Crossed Adjunct-Island, and Crossed RC-Island). Random effects included random intercepts for Subject, Lexicalisation and random slopes for Subject by Session, Subject by Condition, Subject by Structure and Trial by Session (as illustrated above, the experimental tasks were performed over two sessions by every participant). A summary of the optimal model is presented in Table 5-6.

Table 5-6 Summary of the Generalised Additive Model fitted to Difference data. Dependent variable: Difference: RT(gap) – RT (RP).

Reference level: BareWH, Non-Crossed Island. Formula: Difference ~ s(subject, by = session, bs = "re") + s(subject, Condition, bs="re") + s(subject, Structure, bs="re") +s(lexicalisation, bs="re") + s(trial, by=session)+ (Condition * Structure).

Coefficients	Estimate	Std.Er	z value	p-value
Intercept	-41.68	16.66	-2.50	0.012
Condition				
NonCrossed_Island	—			
Crossed_NonIsland	40.14	19.22	2.08	0.036
Crossed_AdIsland	55.45	19.26	2.87	0.003
Crossed_RCIsland	49.30	19.25	2.85	0.010
Structure				
BareWH				
Which	26.18	19.03	1.37	0.168
Condition *Structure				
Crossed_NonIsland * Which	-2.44	27.12	-0.09	0.928
Crossed_AdIsland * Which	-16.44	26.43	-0.62	0.533

Crossed_RCIsland * Which	-19.36	26.43	-0.73	0.463
--------------------------	--------	-------	-------	-------

To facilitate better understanding of the model coefficients, estimates of the model were plotted in the figures below.

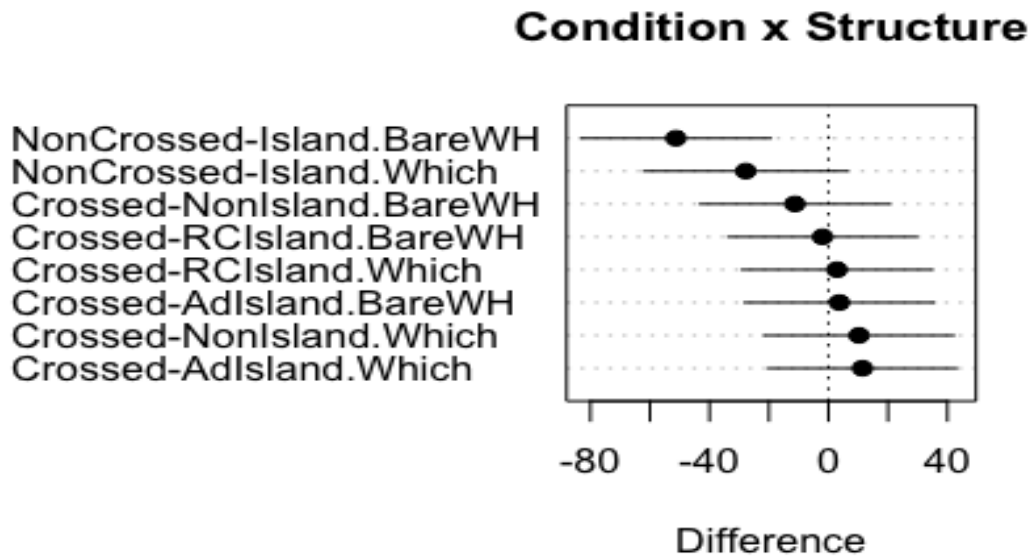


Figure 5-12 The summed effects of the parametric terms in the GAM model.

The summed effects are the predicted response for a certain condition; thus, all partial effects that apply to these conditions are summed up, including the intercept.

From the model and the plot, the statistical analysis revealed no robust main effect for structure (estimate= 26.18, $t= 1.37$, $p= 0.168$). In bare *wh*-questions, there is a significant difference between differential RTs across conditions. The mean of differential RTs is significantly smaller, in comparison to the mean of differential RTs in the non-crossed island, in crossed non-island (estimate=40.14, $t=2.08$, $p=0.036$), crossed RC island (estimate=49.30, $t=2.85$, $p=0.01$) and crossed adjunct island conditions (estimate=55.45, $t=2.87$, $p=0.003$). Furthermore, the effect size is larger in the crossed adjunct island condition. On the whole, the differential RTs between structures with gaps and structures with RPs remain negative in these conditions. However, the RT differential between gap dependencies and RP dependencies becomes smaller for more demanding structures (i.e. when the dependency crosses an island/non-island clause boundary).

As for D-linked questions, there is a statistically significant difference between the means of differential RTs in the non-crossed island and crossed adjunct island conditions (estimate= 39.00, $t = 2.02$, $p = 0.04$), suggesting that RPs are processed faster in the latter than the former. Moreover, there is no robust statistically significant difference between the means of differential RTs in the non-crossed island and crossed RC island conditions (estimate= 29.94, $t = 1.55$, $p = 0.12$), and in the non-crossed island and crossed non-island conditions (estimate= 37.69, $t = 1.88$, $p = 0.06$).

To determine whether RPs are processed faster than gaps in a certain condition, we depend on the plot in Figure 5-12 and the estimate of each condition when it occurs as the intercept of a model. The plot and intercept estimates revealed that, in bare *wh*-questions, gaps are processed significantly faster than RPs in the non-crossed island condition (estimate=-41.68, $t = -2.50$, $p = 0.012$). There is moreover no significant difference between the processing of gap and RP dependencies in the crossed non-island condition (estimate=-1.55, $t = -0.09$, $p = 0.92$), crossed RC-island (estimate= 7.62, $t = 0.45$, $p = 0.65$) and crossed adjunct island condition (estimate= 13.77, $t = 0.81$, $p = 0.41$). This indicates that RPs are never processed significantly faster than gaps in a particular condition in bare *wh*-questions.

In D-linked *wh*-questions, the results revealed no significant difference between the processing of gap dependencies and RP dependencies in the non-crossed island (estimate= -15.50, $t = -0.93$, $p = 0.35$), crossed non-island (estimate=22.19, $t = 1.22$, $p = 0.22$), crossed adjunct island (estimate=23.50, $t = 1.38$, $p = 0.17$) and crossed RC island (estimate= 14.44, $t = 0.85$, $p = 0.39$) conditions. Therefore, there is no significant difference between the means of differential RTs in RP structures and gap structures in the different conditions.

In summary, our results revealed that the only condition in which structures with gaps are robustly faster to process than structures with RPs is the non-crossed-island condition in non-D-linked *wh*-questions. In all other conditions, the CI crosses 0, meaning there is no significant processing speed difference between the two structures (with gap or RP). However, there are interesting numerical trends (with a positive score) in the most complex structures, both in D-linked and non-D-linked *wh*-questions.

5.3.6 Discussion

The current experimental study examined the amelioration effect of RPs in *wh*-questions in Baha Arabic. Experimental sentences include non-D-linked bare *wh*-questions and D-linked *wh*-questions. All experimental sentences were preceded by context sentences to increase their plausibility and thereby optimise the conditions for a facilitation effect to emerge.

Under the hypothesis that intrusive RPs facilitate the processing of complex structures, it was expected that structures with RPs will be processed faster than structures with gaps in more demanding contexts. This phenomenon has been observed in English, where RP dependencies are processed faster than gap dependencies in embedded contexts (Hofmeister and Norcliffe, 2013) and inside island structures (Hammerly, 2019).

The results of the present study reveal that structures with RPs are never processed faster than structures with gaps across the different conditions in both types of questions, but that differences between RTs in gap structures and RP structures decreases as processing demands increase; i.e. RPs are less disruptive in more complex structures.

Starting with non-D-linked bare *wh*-questions, structures with RPs are processed more slowly than structures with gaps in 'non-complex' non-D-linked *wh*-questions. At the same time, however, the size of this difference decreases significantly when a non-island or island clause boundary is crossed. This effect even goes in the other direction, at least numerically. Interestingly, the effect size of this decrease in processing cost of RP dependencies is different for each of the three conditions: the largest effect is observed in adjunct islands, followed by RC island and non-island contexts. In the adjunct and RC island conditions, differential RTs (calculated as RT (gap)-RT (RP)) are positive, suggesting that RPs are processed faster than gaps. However, as illustrated above, this facilitation effect does not reach significance.

As for D-linked *wh*-questions, there was no significant difference between the processing of structures with RPs and those with gaps across the different conditions. This suggests that parsers are not as surprised by the presence of RPs in D-linked *wh*-questions as in non-D-linked *wh*-questions.

Hence, the conclusion we can obtain from the results of the present study is that RPs disrupt rather than facilitate processing in bare *wh*-questions, as they are associated with longer RTs in comparison to gaps. However, this processing cost decreases as

the processing complexity increases, as suggested by difference data. In D-linked *wh*-questions, on the other hand, RPs are not disruptive. It seems that it is difficult for participants to interpret the pronoun as coreferential with the non-D-linked filler phrases, while this is less difficult for D-linked fillers. These results indicate that resumption in Baha Arabic *wh*-questions is sensitive to d-linking, as expected based on Frazier and Clifton (2002). However, it should be noted that RP dependencies are never processed significantly faster than gap dependencies in D-linked *wh*-questions (despite the existence of a numerical trend in that direction).

How can we interpret this pattern of results in light of the processing accounts of RPs?

A dominant hypothesis in the literature on resumption is that dependencies with intrusive RPs are ultimately anaphoric dependencies, such that parsers resolve the FGDs anaphorically rather than syntactically (Erteschik-Shir, 1992; Frazier and Clifton, 2002; Alexopoulou, 2010; Chacón, 2019).

A comprehensive theory in line with this hypothesis is provided by Chacón (2019). Chacón's (2019) processing theory assumes that intrusive RP dependencies are resolved anaphorically. However, he further argued that comprehenders only consider resumption to resolve the FGDs when they have already forgotten the gaps. Upon encountering a filler phrase, the parser creates a representation of the predicate containing a gap, so that the filler phrase will be assigned with its theta-role at the gap position. This representation must be held active in WM throughout the processing of the FGD. However, this representation is subject to a decay effect over time. Processing complex dependencies that span multiple clause boundaries and island clause boundaries can overload WM resources, which further hinders the parser's ability to maintain the representation of gaps active in WM. This, in turn, will lead participants to be less sensitive to whether or not the filler phrase successfully binds a gap. When a pronoun is encountered, participants will initiate a search process for a potential antecedent for the pronoun. If the parser identifies the filler phrase as the antecedent of the pronoun, the filler phrase can be linked to the dependency structure, which in turn makes it possible to obtain a coherent interpretation of the dependency, despite its ungrammaticality.

Looking back at the results of the current study, we found that, in bare *wh*-questions, resumptive pronouns in simple dependencies are disruptive. The cost of processing RP dependencies decreases as the processing demands of a dependency structure increase. Thus, in light of Chacón's (2019) account, we can interpret the cost of processing RPs in non-complex contexts as reflecting a reanalysis effect: participants

expect a gap, but encounter a pronoun. In complex contexts, however, the representation of gaps is lost. Hence, parsers are not surprised when they encounter the pronoun, as they already do not expect to find a gap. Instead, they initiate a search process for a potential antecedent for the pronoun.

If this analysis is on the right track, then the difference between bare *wh*-questions and D-linked *wh*-questions observed in the present study is in line with expectations. In other words, if resumption initiates an anaphoric resolution of the dependency, it is unsurprising that resumption is not disruptive in D-linked *wh*-questions: specifically, if better-established discourse referents make better antecedents, D-linked fillers are better interpreted as coreferent with the pronoun than bare inanimate filler phrases (Frazier and Clifton, 2002; Alexopoulou and Keller, 2013). Parsers will accordingly find it difficult to link the pronoun with 'what' than with 'which-N' phrases, both in complex and non-complex contexts.

However, the current study could not detect any processing facilitation effect of RPs; i.e. RP dependencies are never processed faster than gap dependencies, even in complex contexts, in both types of *wh*-questions. We thus propose the following interpretations for this behaviour in each type of dependency structure.

Starting with bare *wh*-questions, it is possible that RP dependencies and gap dependencies in complex contexts are both complex to process due to the difficulty associated with obtaining a sensible interpretation of the dependency; in short, parsers are unable to assign filler phrases with their theta roles. In gap conditions, dependencies are too complex to resolve inside islands, whether this is due to the application of a grammatical constraint on movement (Wagers and Phillips, 2009), or the low activation level of filler phrases in WM, which in turn leads the mental representation of the filler phrase to be lost from WM (Kluender, 1991) and increases the difficulty of retrieving it at the integration site (Hofmeister and Sag, 2010).

In RP dependencies, similar to gap conditions, parsers are not able to assign fillers with their theta role, as the coreference relation between the pronoun and the filler phrase 'what' is not successfully established. In other words, the *wh*-phrase 'what' is often described as inherently non-discourse linked, in that it cannot be interpreted referentially in the absence of a discourse context in which its referent is explicitly defined. If RPs initiate anaphoric resolution of the *wh*-dependency, parsers will be less likely to consider 'what' as an antecedent for the pronoun.

As explained above, the context sentences did not mention the referent of the pronouns and their associate filler phrases. The filler in question was therefore not discourse-linked (except if inherently so, as in *which*-questions). If intrusive RPs resolve dependencies anaphorically, as indicated by the fact that RPs are not disruptive in D-linked *wh*-questions as they are in bare *wh*-questions, we expect that the lack of referential discourse context makes it difficult for participants to interpret 'what' referentially (and, therefore, link it with the pronoun).

It is also important to note that these structures are likely considered ungrammatical, as indicated by Study 1. Thus, the disruption in processing RP structures in simple non-D-linked *wh*-questions is expected due to the ungrammaticality of RPs in such contexts. The observation that the disruption caused by RPs in non-D-linked *wh*-questions is reduced as processing demands increase might further suggest that parsers become less sensitive to the ungrammaticality of RPs as processing demands increase.

As for D-linked *wh*-questions, we also found no significant difference between the processing of RP dependencies and gap dependencies across conditions; i.e. RPs are never processed faster than gaps. These results suggest that RPs do not facilitate the processing of complex island-violating dependencies.

However, this conclusion might be affected by a confounding factor. Indeed, the literature on D-linked questions argues that D-linked fillers ameliorate islands as resumptive pronouns do. Alexopoulou and Keller (2015) investigated how D-linking and resumption ameliorate island effects in Greek and English. Their findings suggested that while both ameliorate island effects, D-linking has the greatest amelioration effect.

It is therefore possible that the results of the present study reflect the fact that D-linked fillers ameliorate islands in both gap dependencies and RP dependencies, and that the amelioration effect of D-linked fillers exceeds the amelioration effect of RPs. Unfortunately, the design of the current study does not allow us to disentangle the effect of resumption from the effect of D-linking on processing island-violating dependencies.

The conclusion so far is that RPs in Baha Arabic *wh*-questions do not facilitate processing in complex-to-process structures. These structures are never processed significantly faster than gaps, either in D-linked questions or non-D-linked *wh*-questions. However, the results of the current study do give some support to the

hypothesis that intrusive RP dependencies are interpreted anaphorically (as shown by the fact that RPs are disruptive in non-D-linked bare *wh*-questions, but not in D-linked *wh*-questions).

An important point is in order. In the current study, following Alexopoulou (2010), we assumed that RPs in *wh*-questions, both D-linked and non-D-linked, are intrusive (i.e. not grammatically licensed). However, it has been argued that RPs in D-linked *wh*-questions in Arabic are in fact grammatically licensed (Aoun et al., 2009). Unlike the processing accounts that we consider in the current study, syntactic accounts attribute a grammatical function to RPs in D-linked *wh*-questions. In other words, while the processing theories argue that RPs in D-linked *wh*-questions are processed as discourse pronouns, syntactic theories assume that RPs in D-linked *wh*-questions are processed as bound variable pronouns in syntactic binding dependencies.

Our results can be interpreted under this syntactic view as follows: RPs are disruptive in non-D-linked questions, but are not disruptive in D-linked *wh*-questions, because they are grammatically licensed in the latter, but not in the former. Unfortunately, the lack of acceptability judgement data for island-violating D-linked *wh*-questions makes it difficult for us to determine the status of RPs in D-linked *wh*-questions (i.e. whether or not RPs are grammatically licensed in such contexts).

Therefore, the nature of the interaction between the discourse properties of filler phrases and resumption merits further investigation, as does the question of whether this interaction is better accommodated under a syntactic account or a processing account.

5.3.7 Outstanding questions

1. What are the discourse properties of RPs inside islands in Baha Arabic *wh*-questions?
2. Does manipulating the discourse properties of filler phrases affect the processing and acceptability of RP dependencies in Baha Arabic *wh*-questions?

These questions will be addressed in the next chapter.

5.4 Conclusion

Using two self-paced reading tasks, we investigated the effect of resumptive pronouns in the online processing of dependency structures in Baha Arabic. In the first task, we found that RPs in relative clauses did facilitate processing. This result is compatible with the results of the acceptability experiment, which suggested that RPs are both obligatory and grammatically licensed in relative clauses.

No facilitation effect of RPs in bare *wh*-questions was observed, either when these were non-contextualised (first self-paced reading experiment) or contextualised (second self-paced reading experiment). Instead, we found that RPs increase processing time in non-complex contexts. However, the cost of processing RP dependencies decreases as processing demands increase. In D-linked *wh*-questions, on the other hand, we found that RPs do not increase processing time in non-complex contexts. However, RPs are never processed significantly faster than gaps, even inside islands; hence, the evidence so far is that RPs do not significantly facilitate the processing of islands in Baha Arabic *wh*-questions.

These results are consistent with an anaphoric analysis of intrusive RPs (Erteschik-Shir, 1992; Frazier and Clifton, 2002; Alexopoulou, 2010; Chacón, 2019). In particular, we assume that the absence of the processing advantage of RPs in bare *wh*-questions is due to the difficulty of interpreting non-D-linked *wh*-phrases as co-referential with the pronoun in absence of a discourse context that licenses a d-linking interpretation for such fillers. The absence of a facilitation effect for RPs in D-linked *wh*-questions could be due to the fact that the amelioration of d-linking on islands exceeds the amelioration effect of RPs.

Nevertheless, our results can also be accommodated under a syntactic theory of d-linking and resumption: specifically, RPs are disruptive in non-D-linked questions but are not disruptive in D-linked questions because they are grammatically licensed in the latter, but not in the former.

In the next chapter, we will investigate the extent to which resumption interacts with the discourse properties of filler phrases. We will further examine whether, if such interaction is present, it is better accommodated under a syntactic or a processing account. In other words, the following chapter will explore whether RPs in D-linked *wh*-questions in Baha Arabic are processed as discourse pronouns or as syntactic variables in syntactic binding dependencies.

Chapter 6 Discourse Linking and Resumption

6.1 Introduction

The experimental studies reported in previous chapters support the distinction between ‘true’ or grammatical RPs and intrusive RPs in Baha Arabic. Specifically, RPs are grammatically licensed in relative clauses, but are not licensed by the grammar in *wh*-questions. We accordingly argue that relative clauses are derived by binding relations, a conclusion supported by the high acceptability of island-violating relative clauses with RPs. On the other hand, we argue that bare *wh*-questions are movement dependencies, a conclusion supported by the low acceptability of island-violating bare *wh*-questions with or without RPs.

Following Alexopoulou (2010), we hypothesised that true RPs in relative clauses are required for grammatical purposes, while RPs in *wh*-interrogatives are required for processing reasons. However, the experimental studies reported in previous chapters did not detect any processing advantage for RPs in interrogatives, either in bare (non-discourse-linked, or ‘non-D-linked’) *wh*-questions or in D-linked *wh*-questions. Instead, results revealed that RPs are disruptive in non-complex non-D-linked *wh*-questions, but not in non-complex D-linked *wh*-questions. This is consistent with the hypothesis that intrusive RPs can license an anaphoric interpretation of Filler-Gap Dependencies (FGDs) under certain circumstances (Alexopoulou, 2010; Chacón, 2019). Such circumstances include dependencies where filler phrases can be referentially interpreted.

However, if this is the case, why do RPs confer no processing advantage in D-linked *wh*-questions, where the filler phrase is referential? Our contention is that this result might be due to a confound in our design, which did not allow us to disentangle the processing advantage of D-linking from the processing advantage of RPs in islands. In other words, we speculate that the processing advantage of D-linking in complex FGDs exceeds the processing advantage of RPs, which in turn makes it difficult to observe the processing effect of RPs.

The current study accordingly aims to investigate the nature of the interaction between discourse properties of filler phrases and resumption; that is, we aim to test the extent to which manipulating the discourse properties of filler phrases affects the processing and acceptability of RP-dependencies in island-violating *wh*-questions in Baha Arabic. A further aim of this experiment is to disentangle the impact of D-linking from the impact of resumption on islands in island-violating *wh*-questions in Baha Arabic.

The remainder of this chapter is organised as follows. Section 6.2 briefly reviews the claims that D-linking properties of filler phrases increase RP acceptance across languages. Section 6.3 briefly reviews the syntactic and psycholinguistic accounts of how the D-linking properties of filler phrases affect FGDs more generally. Section 6.3.1 presents the syntactic account of the interaction between D-linking and resumption, while Section 6.3.2 presents the psycholinguistic account of the interaction between D-linking and resumption. Section 6.3.3 discusses how the syntactic and psycholinguistic accounts differ in terms of how they predict the interaction between D-linking and resumption. Section 6.4 presents a different account of the processing advantage of *which*-questions relative to bare *what*-questions (compared to the D-linking account). Section 6.6 outlines our research questions, hypotheses and predictions. Section 6.6.1 presents the methods used in the present experiment, while Section 6.6.2 presents the results and a discussion of the data. Section 6.7 summarises and discusses the findings. Finally, Section 6.8 concludes the chapter.

6.2 Discourse-linking and resumption

Cross-linguistically, RPs are rarely accepted in *wh*-questions unless the filler phrase is D-linked, either inherently (i.e. *which*-N) or when preceded by a specific context (Abdel Razaq, 2011). Examples from Hebrew, Albanian, Modern Standard Arabic (MSA), Lebanese Arabic and Jordanian Arabic are provided below.

1. a. eyze student nifgaSta (ito)
 which student you-met with-him
 ‘Which student did you meet?’
- b. *mi nifgaSta ito
 who you-met with-him
 ‘Who did you meet with?’
 (Hebrew, Sharvit, 1999)
2. a. Çfarë (*e) solli Ana?
 What 3S.ACC brought Ana.NOM
 ‘What did Ana bring?’
- b. Cil-in libër (e) solli Ana?

- which-the.ACC book 3S.ACC brought Ana.NOM
 ‘Which book did Ana bring?’
 (Albanian, Kallulli, 2009)
3. a. maaðaaɪ ʔiʃtarat (*-hui) laila min al-maktabati?
 what_i bought-3FS (*-it_i) Laila from the-bookstore
 ‘What_i did Laila buy (*it_i) from the bookstore?’
- b. ʔayya kitaabini ʔiʃtarat(hu_i) laila min
 which book_i bought-3FS(-it_i) Laila from
 al-maktabati?
 the-bookstore
 ‘Which book_i did Laila buy (it_i) from the bookstore?’
 (MSA, Tucker et al., 2019)
4. a. šu štarayt-i(*-h) b-l-maktabe ?
 what_i bought-2FS-it_i from-the-bookstore
 ‘What_i did you buy (*it_i) from the bookstore?’
- b. ayya mmasil šeft-o b-l-matʕam?
 which actor saw-2MS-him in-the-restaurant
 ‘Which actor did you see in the restaurant?’
 (Lebanese Arabic, Aoun et al., 2009, p.128)
5. Context: Qnde kull l-ʔawaaʕi l-jadeede
 have.1S all the-clothes the-new
 ‘I’ve got all the new clothes’
- a. ʔeyš tjarribt b-l-ʔawaaʕi?
 what try.2FS in-the-first
 ‘What did you try first?’
- b. ʔeyš tjarribt-i b- l-ʔawaaʕi?
 what try.2FS-it in-the-first

'What did you try first?'

(Jordanian Arabic, Abdel Razaq, 2011, pp. 162-163)

However, the exact mechanism that accounts for the interaction between resumption and D-linking differs according to one's perspective on the nature of the D-linking effect. As will be illustrated below, the discourse properties of filler phrases are known to affect both the processing and acceptability of island-violating FGDs (Cinque, 1990; Szabolcsi and Zwarts, 1993; Rizzi, 2001; Rizzi, 2004; Goodall, 2017). Therefore, in order to understand the nature of the interaction between resumption and D-linking, we first need to illustrate how the D-linking properties of filler phrases affect the acceptability and processing of island-violating dependencies, as well as how syntacticians and psycholinguists account for this effect. Accordingly, this will be the topic of the next section.

6.3 Nature of the D-linking effect

It has been argued that island effects are ameliorated, or even repaired, when replacing a bare *wh*-phrase (e.g. *what*) with a *which-N* phrase, as in (6) below (Maling and Zaenen, 1982; Cinque, 1990; Rizzi, 1990; De Swart, 1992; Chung, 1994):

6. a. ***What**_i do you believe [the claim that the man bought ____i]
 b. ?? **Which car**_i do you believe [the claim that the man bought ____i]?

(Goodall, 2014, p.1)

This effect has been demonstrated in a number of experimental studies. For example, Hofmeister and Sag (2010) tested the effect of *which-N* phrases on island-violating dependencies (see 7 below), using an online self-paced reading task and an offline acceptability judgment task. Results showed that replacing a bare *wh*-phrase with a *which-N* phrase improved acceptability ratings ($F1(1,20) = 48.741, p < 0.0001$; $F2(1,35) = 39.494, p < 0.0001$); in the RT data, moreover, the main effect of filler phrase complexity was highly significant when averaging RTs over the complementiser and spillover regions (bolded in (7)) ($F1(1,24) = 18.365, p < 0.001$; $F2(1,35) = 22.723, p < 0.001$):

7. I saw (who | which convict) Emma doubted the report **that we had captured ___ in the** nationwide FBI manhunt.

The effect of *which-N* phrases on FGDs is also found in multiple *wh*-questions. To illustrate, in English multiple *wh*-questions, one of the *wh*-phrases must appear at the

left edge of the clause. The more syntactically prominent *wh*-phrase is preferred to be fronted, as illustrated in (8a) and (8b):

8. a. I wonder who bought what.
b. *I wonder what who bought.

(Goodall, 2017)

This phenomenon is known as the ‘superiority effect’ (Chomsky, 1973). However, it is also claimed that this effect is ameliorated or even erased when the *wh*-phrases are *which-N* phrases (Pesetsky, 1987), as in (9).

9. a. I wonder which man bought which car.
b. I wonder which car which man bought.

(Goodall, 2017)

The nature of this processing effect is still largely controversial. A dominant hypothesis attributes the effect of *which-N* phrases on island-violating dependencies and multiple *wh*-questions to the ‘D-linked’ (Pesetsky, 1987) or ‘referential’ (Rizzi, 1990; Cinque, 1990) nature of filler phrases.

Pesetsky (1987) suggested that *which-N* phrases are D-linked, in the sense that the set of possible felicitous answers to *which*-questions is limited to members of a set of contextually salient entities that are known to both the speaker and the hearer. For example, in (10a), the range of possible answers is limited to a set of books, assumed to be present in the common ground of both speaker and hearer. By contrast, no such restriction is implied with questions created with bare *wh*-words; in these cases (such as (10b), the range of possible answers can include any item that is readable.

10. a. Which book_j do you want to read ___i?
b. What_j do you want to read ___i?

Note that *wh*-phrases might be inherently D-linked (e.g., ‘which book’), or contextually D-linked (i.e. without D-linking marking, like ‘what’ or ‘who’) when presented in a discourse context and interpreted as linked to a previous discourse referent. The following examples from (Pesetsky, 1987, p.309) illustrate this point.

11. a. I know what just about everybody was asked to do, but what did who (actually) do?

- b. I know that we need to install transistor A, transistor B, and transistor C, and I know that these three holes are for transistors, but I'll be damned if I can figure out from the instructions where what goes!

In the above examples, there is no overt D-linking mark on wh-phrases. However, the discourse context limits the reference of wh-phrases to what is already known by the interlocutors. Hence, the wh-phrases are interpreted as D-linked, and the superiority effect is argued to be absent. If the wh-phrases are interpreted as non-D-linked, such questions would constitute a violation of superiority effects. When such questions occur out of the blue, as in (12), they are ungrammatical.

12. a. *Guess what did who do (what)
b. *Guess where what goes (where)

(Abdel Razaq, 2011)

The wh-phrases in (11a-b) and (12a-b) differ in terms of their discourse properties: specifically, the wh-phrases are D-linked in (11a-b), but are not so in (12a-b). Hence, the non-D-linked version is ruled out.

The mechanisms by which D-linking affects the acceptability of FGDs remain controversial. More specifically, some argue that the discourse properties of filler phrases affect the underlying syntactic structure of FGDs (Pesetsky, 1987), while others argue that these discourse properties merely have a facilitation effect (Frazier and Clifton, 2002).

These hypotheses have consequences regarding how the interaction between resumption and D-linking is accounted for. Those who adopt the former hypothesis assume that RPs are required for grammatical purposes in binding dependencies; RPs are either obligatory overt, as in Arabic varieties (Aoun et al., 2009), or covert, as in English (Kallulli, 2009).³ On the other hand, those who adopt the latter hypothesis assume that RPs are intrusive and used to facilitate the processing of

³ There are some accounts that assume that the movement of the wh-phrase in D-linked questions is not in fact wh-movement. For instance, Gad (2011) argued that D-linked questions and cleft wh-questions are derived through focus movement, rather than wh-movement, in Egyptian Arabic. Boeckx and Grohmann (2004) further proposed that D-linked questions are derived through the application of a sub-move, where D is stranded and later realised as a resumptive pronoun. The extracted materials move to the left periphery of the clause, which expresses discourse properties such as presuppositionality.

ungrammatical FGDs in line with D-linked fillers. In other words, resumption and D-linking are given a syntactic interpretation under the former account, but a processing interpretation under the second account. These two accounts are presented in more detail in the next section.

6.3.1 D-linking and resumption as syntactic phenomena

The distributional pattern of RPs in Arabic interrogatives has been largely investigated in terms of formal syntax. The dominant hypothesis is that *wh*-phrases that are allowed to occur with resumptive pronouns must be D-linked, either inherently or contextually.

Aoun et al. (2009) find support for this hypothesis in MSA and Lebanese Arabic data. These authors argue that, in both of these languages, RPs are allowed with inherently D-linked fillers (i.e. *which-N*). As for non-inherently D-linked fillers, Aoun et al. (2009) found that RPs in the two languages are allowed with ‘who’ but not with ‘what’. They attributed this difference to the fact that the filler phrase ‘what’, in both MSA and Lebanese Arabic, cannot be D-linked even in the presence of a licensing discourse context. See (13) and (14) below.

13. * šu štarayt-i
what bought.2FS-it
‘What did you buy?’

(Lebanese Arabic, Abdel Razaq, 2011)

14. maḡḡaai ʔištara(*-hu_i) laila min al-maktabati?
what_i bought.3FS(*-it_i) Laila from the-bookstore
“What_i did Laila buy (*it_i) from the bookstore?”

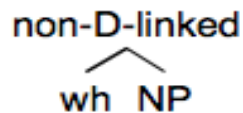
(MSA, Tucker et al., 2019)

Hence, these authors argued that, in Lebanese Arabic and MSA, only ‘who’ can be contextually d-linked, while ‘what’ cannot; hence, RPs are allowed with ‘*which-N*’ and ‘who’, but not with ‘what’ (Aoun and Choueiri, 1999; Aoun et al., 2009).

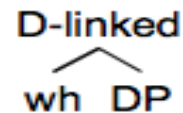
Based on these observations, Aoun and Choueiri (1999) and Aoun et al. (2009) go on to suggest that those *wh*-phrases that can be D-linked and therefore bind a resumptive pronoun comprise both a *wh*-element that bears the ‘*wh*’ feature and a full DP. On the other hand, *wh*-phrases that cannot be D-linked, and therefore cannot

bind a resumptive pronoun, are composed of a wh-element that bears the ‘wh’ feature and an NP. More specifically, Aoun et al. (2009) suggest that inherently and contextually D-linked wh-fillers (i.e. the wh-words *miin/man* (‘who’) and *ʔayy(a)* (‘which’)) have the representation shown in (15b), whereas the non-d-linked fillers (i.e. the wh-expression *šu/maadaai* (‘what’)) has the representation shown in (15a).

15. a.



b.



Abdel Razaq (2011) reports data from Jordanian Arabic that supports this conclusion. According to Abdel Razaq (2011), Jordanian Arabic uses two forms of ‘what’: the first, *šū*, is similar to the one used in Lebanese Arabic, while the other, *ʔayš*, is similar to the one used in most Arabic varieties (and is also the same as that used in Baha Arabic). He further provides data showing that *šū* (the Lebanese form of ‘what’) cannot be D-linked when preceded by a context in Jordanian Arabic, and hence does not allow resumptive pronouns. On the other hand, the other form of ‘what’ (*ʔayš*) can be D-linked and can henceforth bind a RP. See (16).

16. Context: ʕnde kull l-ʔawaaʕi l-jadeede
 have.1S all the-clothes the-new

‘I’ve got all the new clothes’

a. ʔayš tjarribt b-l-ʔawwal?
 what try.2FS in-the-first
 ‘What did you try first?’

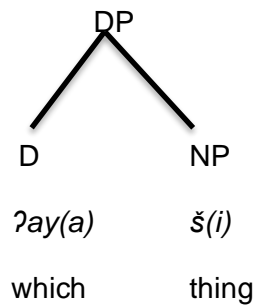
b. ʔayš tjarribt-i b-l-ʔawwal?
 what try.2FS-it in-the-first
 ‘What did you try first?’

c. *šū tjarribt-i b-l-ʔawwal?
 what try.2FS-it in-the-first
 ‘What did you try first?’

To account for this difference, Abdel Razaq goes on to argue that the wh-expression *ʔayš* ('what') is a syntactic variant of *ʔayy(a) š(i)* ('which thing'); he argued that the wh-form *ʔayš* ('what') is bi-morphemic that is derived from the combination of the wh-element *ʔay(a)* ('which'), and the indefinite (NP) *š(i)* ('thing'). The two-word wh-phrase *ʔay(a) š(i)* undergoes phrasal spellout as a single chunk, yielding the form *ʔeyš*.

Abdel Razaq (2011) argued that the two-word wh-phrase *ʔay(a) š(i)* ('which thing') is a DP similar to its English counterpart 'which thing' (Chomsky, 1995). Inside the DP structure, the wh-element *ʔay(a)* ('which') occupies the D position and takes the NP *š(i)* ('thing') as its complement. See (17).

17.



In Arabic, it is commonly argued that island-violating D-linked wh-questions are derived by binding relations rather than movement relations. This argument is based on the observation that island-violating D-linked questions obligatorily occur with RPs, and are insensitive to island effects (Aoun et al., 2009).⁴

Furthermore, there has been no investigation on the independent effect of D-linking on islands. The common view is that island-violating D-linked questions with gaps are ungrammatical, as the binding dependency cannot be established in the absence of overt RPs; in other words, there is no argument that D-linked fillers weaken island effects in the absence of RPs. The following notation conventions for acceptability are used to illustrate the acceptability of island-violating D-linked questions with and without RPs in Arabic varieties.

18. a. ya: Su:ra.F li-bni-ha Samer ysʔal ida

⁴ Non-referential filler phrases (e.g. the wh-expression *š(u)/maatha* ('what') in Lebanese Arabic and MSA) always undergo overt movement to Spec,CP of the wh-interrogative. This is manifested in their sensitivity to islands, even in the presence of RPs (Aoun et al., 2009).

which picture.F of-son-her Samer ask.3MS if
kull mraya jagagat-ha
every woman tore.3FS-3F

'Which photo of her son did Samer wonder if every woman tore [it]?

b. * ya: Su:ra.F li-bni-ha Samer ysʔal ida
which picture.F of-son-her Samer ask.3MS if
kull mraya šagagat__
every woman tore__

'Which photo of her son did Samer wonder if every woman tore _ ?

(Iraqi Arabic, Sterian, 2016)

In summary, the following hypotheses have been proposed for Arabic:

1. WH-questions that allow RPs must be D-linked, either inherently or contextually (Aoun and Choueiri, 1999; Aoun et al., 2009).
2. WH-dependencies with resumptive pronouns are insensitive to islands, and therefore do not involve movement (Aoun et al., 2009).
3. D-linked fillers do not weaken island effects in the absence of RPs.
4. The dominant view adopted by Arabic syntacticians is that RPs are only allowed when they are grammatically licensed, either in D-linked questions or relative clauses (Aoun et al., 2009). In other words, Arabic does not feature intrusive RPs.⁵

6.3.2 D-linking and resumption as processing phenomena

Another dominant hypothesis regarding the amelioration effect of D-linking argues that D-linking does not alter the underlying syntactic structure of FGDs, but rather simply facilitates the processing of complex-to-process FGDs; in other words, the effect of D-linking is argued to be extra-grammatical (Frazier and Clifton, 2002;

⁵ Sterian (2016) claimed that RPs inside island structures in Arabic wh-dependencies are intrusive. This claim was based on the observation that RPs do not always save all types of islands across the different varieties of Arabic.

Hofmeister and Sag, 2010; Goodall, 2014; Goodall, 2017). To illustrate, Frazier and Clifton (2002) argued that upon encountering *which-N* phrases, the parser immediately assigns them a discourse representation along with a syntactic representation. When the dependency cannot be established in syntax, the parser resorts to the discourse representation of the filler phrase to resolve the dependency anaphorically; this allows parsers to obtain a sensible interpretation of the dependency regardless of its grammatical status. Presumably, this means that non-D-linked *wh*-fillers, unlike D-linked *wh*-fillers, are interpreted in syntactic representation but not in discourse representation.

This hypothesis is supported by the observation that the amelioration effect of D-linked fillers has been observed in non-island and island contexts alike in acceptability judgment data (Goodall, 2014) and in reading time data (Hofmeister, 2007; Hofmeister, 2011). For example, Goodall (2014) tested the effect of filler phrase type on the acceptability of FGDs crossing an island and a non-island clause boundary. Goodall (2014) argued that syntactic accounts of D-linking predict that the D-linking effect is restricted to island-violating dependencies, as these involve violations of grammatical constraints on movement. Processing accounts of D-linking, on the other hand, predict that D-linking should improve the acceptability of FGDs, regardless of whether an embedded island or a non-island clause boundary is crossed.

Goodall (2014) examined these predictions in an acceptability judgment experiment. More specifically, he manipulated the type of the *wh*-filler (bare vs. complex [D-linked]) and the type of the structure in which the gap was located (complex NP vs. *wh*-clause vs. *that*-clause). Some sample experimental sentences are presented below.

19. a. Ungrammatical / Complex NP Island
*What / ***Which of the cars** do you believe the claim that he might buy ___?
- b. Ungrammatical / Wh-island
*What / ***Which of the cars** do you wonder who might buy ___?
- c. Grammatical / Non-island
What / **Which of the cars** do you believe that he might buy ___?

The results showed that dependencies created with complex (D-linked) *wh*-fillers are rated higher than dependencies created with bare *wh*-fillers, regardless of context, and that this increase in acceptability is also uniform in both island and non-island

contexts. Furthermore, island-violating dependencies are rated lower than their non-island counterparts in both bare *wh*-questions and D-linked *wh*-questions. Hence, Goodall (2014) argued that the amelioration effect of D-linked fillers can be better attributed to processing factors that contribute to facilitating the processing complexity of FGDs in general. Notably, the observation that this amelioration effect is uniform in island and non-island contexts is problematic for purely syntactic accounts, to the extent that these accounts attribute the alleviation of island effects to the absence of syntactic movement. Furthermore, Goodall (2014) argued that a distinction between movement and base-generation dependencies should be maintained to account for island effects. In other words, Goodall (2014) argued that movement is involved in the derivation of D-linked and non-D-linked dependencies, but that D-linked filler-phrases ameliorate islands as they facilitate comprehension.

Proponents of the processing theory of D-linking take the fact that resumption is sensitive to D-linking as evidence in support of their account. To illustrate, Frazier and Clifton (2002) found that whether-islands with resumptive pronouns receive higher acceptability scores when the *wh*-phrase is D-linked (20a) than when it is not (20b).

20. a. (*) Which students did the teacher wonder if they had gone to the library?
- b. (*) Who did the teacher wonder if they had gone to the library?

(Frazier and Clifton, 2002)

To account for this result, Frazier and Clifton (2002) argued that ‘the resumptive pronoun is not treated as resumptive at all, but simply as discourse-referential’, and that ‘a D-linked *wh*-phrase sets up a discourse entity to which a pronoun can refer’ (p. 28); while the two questions in (20) are not grammatical, the presence of a D-linked filler and a pronoun aid parsers in arriving at a sensible interpretation of the dependency. This theory adopts the assumption that RPs are processed as discourse pronouns and prefer to find their antecedent in the discourse representation (Sells, 1984).

In summary, processing accounts of intrusive RPs argue that intrusive RPs are processed as normal discourse pronouns rather than bound variables; rather than establishing the dependency between the filler phrase and its related integration position in syntax, it is established in discourse via an anaphoric relation (Prince, 1990; Alexopoulou, 2010; Polinsky et al., 2013; Beltrama and Xiang, 2016; Chacón, 2019).

In previous chapters, two comprehensive theories were presented (Alexopoulou's (2010) and Chacón's (2019)) of how intrusive RPs facilitate processing. Importantly, both theories are based primarily on the hypothesis that the processing advantage of intrusive RPs stems from these pronouns being processed as discourse pronouns, which allows parsers to initiate an anaphoric interpretation for the FGD when resolving such dependencies is difficult or impossible.

6.3.3 Differential predictions

In a nutshell, the syntactic D-linking theory argues that RPs establish an A-bar binding relation with D-linked fillers, while the processing theory argues that RPs establish a co-referential relation with the D-linked fillers. Both of these theories share the prediction that RPs are preferred to gaps inside islands in D-linked wh-questions. However, they make different predictions regarding the magnitude of the amelioration effect, as well as the impact of the salience of discourse referents. The syntactic theory predicts that RPs will fully improve the acceptability of island-violating D-linked questions, while the absence of RPs will render these dependencies completely unacceptable. The processing theory, on the other hand, predicts that RPs will partially improve acceptability of D-linked questions that violate island constraints.

In addition, the processing (discourse-based) theory and the purely syntactic theory make different predictions concerning the sensitivity of resumption to discourse salience. More specifically, the processing (discourse-based) account predicts that the acceptability of RPs is sensitive to the prominence of referents in discourse, whereas purely syntactic accounts make no such prediction. The most substantial difference between the two theories is thus that the former assumes an anaphoric dependency between the filler and the referential discourse pronoun, while the latter assumes a syntactic filler-RP dependency between the filler and bound variable pronoun.

In anaphoric dependencies, the discourse pronoun and its antecedent refer to the same individual in the discourse model (Kush, Lidz, et al., 2015). Accordingly, resolving these anaphoric dependencies is mediated by discourse-based mechanisms. It is argued that the cognitive prominence of a discourse referent in memory aids in resolving anaphoric dependencies when a pronoun is encountered (Cowles et al., 2007); more specifically, when parsing anaphoric dependencies, parsers tend to look for a salient, highly accessible, discourse referent in memory as an antecedent for the pronoun (Clifton and Ferreira, 1987; Ariel, 1990; Cowles et al.,

2007). Thus, the salience of the antecedent's referent in discourse is expected to facilitate the processing of RPs if RPs are processed as discourse pronouns in anaphoric dependencies.

In syntactic-binding dependencies, on the other hand, the pronoun is obligatorily bound by an element in an A-bar position and must be interpreted as a variable bound by an operator at LF. Bound variable pronouns do not refer to a single individual in the discourse model. Instead, their interpretation co-varies with the interpretation of the quantified phrase (Kush, Lidz, et al., 2015); i.e., parsers do not consult the discourse representation when resolving operator bound-variable dependencies (Frazier and Clifton, 2000). Discourse-prominence is accordingly not expected to have an impact on the acceptability and processing of FGDs from the perspective of the syntactic-based theory. Thus, the salience of a referent in discourse is expected to facilitate the processing of RPs only if they are processed as discourse pronouns in anaphoric dependencies, but not when they are processed as bound pronominal variables in syntactic-binding dependencies.

In summary, the syntactic and processing theories both predict that RPs will be preferred to gaps in island-violating D-linked *wh*-questions, but differ in their predictions in relation to:

- the magnitude of the amelioration effect: full improvement effect (syntactic account) vs partial improvement effect (processing-based account).
- the sensitivity to antecedents' salience in discourse: only the processing-based theory predicts that RPs are sensitive to the salience of the referent in discourse.

6.4 Complexity account of *which*-N phrases

The processing effect of *which*-N phrases has also been attributed to factors other than discourse linking. In particular, several theories have been proposed in the literature arguing that D-linking is not the reason why *which*-N phrases facilitate FGDs. For instance, Hofmeister and Sag (2010) argued that the distinguishing property of *which*-N phrases is that these linguistic expressions contain richer structural and semantic content in comparison to bare *wh*-phrases; for example, 'which of your tea cups' is more complex (in terms of semantic and structural content) than 'which cup', which is in turn more informative than 'which one' or 'what'. Semantically and structurally complex *wh*-phrases are more costly to process; however, this processing

cost translates into an increased activation level of the filler in memory, which in turn makes it more accessible for retrieval at the integration site.

Importantly, Hofmeister et al. (2007) argued that *wh*-fillers are not anaphoric, although their interpretation may derive in part from preceding discourse. Thus, Hofmeister et al. (2007) argued that, unlike referential NPs, the activation level of *wh*-fillers in memory is only sensitive to the semantic and structural complexity of the *wh*-filler, but not to the salience of the referent in discourse.

This hypothesis entails that *which*-N phrases will have an advantage in FGD processing, regardless of the status of their discourse-linking properties. Their island sensitivity should be determined by their semantic and syntactic complexity alone. To test this hypothesis, Hofmeister and Sag (2010) tested the island sensitivity of non-referential adjunct filler phrases that varied in terms of their syntactic and semantic complexity, as illustrated in (21a) and (21b) below.

21. a. Julie discerned that the survivor had managed to stay alive for eight days after the crash in the harsh conditions.

SIMPLE: **How long** did Julie observe whether the passenger had survived in the unbelievably harsh conditions?

COMPLEX: **For what period of time after the crash** did Julie observe whether the passenger had survived in the unbelievably harsh conditions?

BASELINE: **How long** did Julie observe that the passenger had survived in the unbelievably harsh conditions?

- b. Andrew overheard the daycare staff discussing how they wanted to get away from the children for a few hours.

SIMPLE: **How long** did Andrew hear whether the children had played during the daycare's afternoon recess?

COMPLEX: **How many hours** did Andrew hear whether the children had played during the daycare's afternoon recess?

BASELINE: **How long** did Andrew hear that the children had played during the daycare's afternoon recess?

Hofmeister and Sag's (2010) account predicts that structurally and semantically complex adjunct fillers will facilitate the processing of island-violating sentences, even though they are not D-linked (as they are non-referential). Results demonstrate that increased syntactic and semantic complexity in adjunct filler phrases does facilitate the processing of FGDs: reading times were significantly longer when the filler phrase was a simple *wh*-word compared to when it was a complex filler phrase ($t_1(27) = 3.484, p < 0.01$; $t_2(23) = 3.513, p < 0.001$).

Other studies have investigated the processing of inherently D-linked fillers (i.e. *which-N*) and contextually D-linked fillers (i.e. 'who' or 'what'). The discourse-linking theory predicts that the cognitive effort involved in interpreting the two types of *wh*-phrases is equivalent, as the set of plausible answers is the same due to their salience in context (Donkers et al., 2013). However, this prediction was not confirmed by experimental studies, such as the work of Avrutin (2000) and Donkers et al. (2013), investigating the processing of (contextually D-linked) bare *wh*-questions vs. (inherently D-linked) *which*-questions.

Avrutin (2000) examined the difference between *which*-questions and *who*-questions in a comprehension study with children and aphasic adult individuals. He used visual contexts in which an explicit set of alternatives was introduced. In each context, there were three animate entities, two of which were of the same type (for example, two horses and one giraffe), as in (22).

22. Scenario: A white horse chases a giraffe that in turn chases a black horse

Subject-initial: Who/Which horse ___ chased the giraffe?

Object-initial: Who/Which horse did the giraffe chase ___?

(Donkers et al., 2013)

Donkers et al.'s (2013) experimental study, on the other hand, investigated healthy adults. This research addressed the question of whether there is a difference between the processing of *which*-questions, generic *which*-person questions, and *who*-questions in Dutch when presented in a felicitous context where an explicit set of alternatives are introduced. Context sentences were constructed in such a way that there are two contrasting members of the same sort and a single member of a different sort; all three were members of a super-ordinate group. An example is given in (23).

23. Context:

Terwijl de dronken bediende een dutje deed, zocht de keizer de nuchtere bediende in de kelder

While the drunken servant took a nap, looked-for the emperor-SUB the sober servant-OBJ in the cellar

- a. Wie heeft de keizer gezocht in
Who-OBJ has the emperor looked-for in
de kelder?
the cellar?
- b. Welke person heeft de keizer gezocht in
Which one-OBJ has the emperor looked-for in
de kelder?
the cellar?
- c. Welke bediende heeft de keizer gezocht in
Which servant-OBJ has the emperor looked-for in
de kelder?
the cellar?

In spite of the licensing of a discourse-linked interpretation of the *wh*-phrases (afforded by the context), both studies still found a difference in the processing of each type of *wh*-phrase. More specifically, Avrutin (2000) found that while no comprehension difficulties was observed with *who*-questions, *which*-questions were comprehended at chance level (in object chains); moreover, Donkers et al. (2013) found that *which*-questions took longer to process at the integration site in comparison to questions created with '*which*-person' and '*who*' (in object chains). Hence, Donkers et al. (2013) argued that it is lexical restriction, rather than D-linking, that is responsible for the effect of *which*-N phrases on processing FGDs.⁶

6 Avrutin's (2000) and Donkers et al.'s (2013) experimental studies further investigated whether the effect of filler type ('*which*' vs '*who*') is affected by the chain type (subject vs object) of the FGD.

Shapiro (2000) investigated the processing of *which*- and *who*-questions using a cross-modal lexical priming study where the referent of the filler phrase was explicitly mentioned in context, as in (24). The square-bracketed numbers in the below example indicate the regions where the priming effect was measured.

24. The soldier is pushing the *unruly student* violently into the street.

Who/Which student [1] is the soldier [2] pushing [3] violently [4] in the street?

Shapiro (2000) assumed that in order to understand discourse, parsers must connect the gap position to the *wh*-word, as well as connect the *wh*-word to the referent that is explicitly mentioned in the previous context sentence (i.e. 'unruly student'); these three elements thus create an 'interpretive chain'. Shapiro (2000) found that a priming effect was observed at the *wh*-word and integration site in both types of *wh*-questions (near significant result, in case of *which*-N at the integration site), and at the post-gap position in *which*-questions. Despite the fact that both types of questions are processed referentially, as indicated by the priming effect at *wh*-words suggesting that the potential referent is active at the *wh*-word, there are still differences in the processing of both types of questions. This observation casts doubt on a purely D-linking-based explanation for the contrast between *which*-N phrases and other *wh*-phrases in terms of island sensitivity.

In summary, if D-linking does not explain the contrast between island-violating bare *wh*-questions and island-violating *which*-questions, these two types of questions are predicted to interact in different ways with island effects, even if both are to some extent interpreted referentially.

6.5 Working memory capacity and RPs

As in the previous experiments, we will test how resumption interacts with individual variation in working memory (WM) capacities. It has been proposed that RPs are only required in cases where the parser is unable to hold the filler phrase active in WM due to processing overload (Ariel, 1999; Hofmeister and Norcliffe, 2013; Hammerly, 2019; Chacón, 2019). Accordingly, we will assume that WM capacity is correlated with the ability to hold filler phrases active in long-distance dependencies and across island structures. Reliance on RPs in such structures might therefore be greater for individuals with lower WM capacity.

6.6 Study 4: Research questions and hypotheses

The aims of this study are as follows: (i) to ascertain the extent to which RPs can alleviate island violations in Baha Arabic wh-questions; (ii) to investigate the sensitivity of island-internal RPs to discourse salience in Baha Arabic; (iii) to understand the source of the amelioration effect of *which*-N phrases, if present, on island-violating dependencies. Our research questions are listed below:

1. Do RPs ameliorate island effects in contextually and inherently D-linked wh-questions?
2. Are RPs in island-violating structures subject to discourse licensing, independent of the properties of the filler?
3. Do *which*-N fillers improve island-violating dependencies with gaps compared to what fillers?

These questions will be examined in two self-paced experiments: one using acceptability judgement, the other a comprehension question. The two self-paced reading sub-experiments are used to ascertain whether the type of task accompanying the self-paced reading might have a confounding effect in previous studies, which has tended to use comprehension question as the follow-up task.

The present study focuses exclusively on island-violating dependencies for two main reasons:

1. Island-violating dependencies constitute an environment that is 'hostile' to wh-movement, making it optimal for clarifying the extent to which RPs are syntactic variables or discourse pronouns.
2. Island-violating dependencies represent the most typical example of complex-to-process FGDs and are therefore an ideal environment for testing the effect of *which*-N phrases on the processing and acceptability of FGDs.

RC-islands were excluded in the current study; we focused solely on adjunct islands. This is because adjunct islands, although often described as strong islands, have also been argued to be weaker (Cinque, 1990; Truswell, 2007), and were also found to be sensitive to extra-grammatical factors, such as satiation effects (Chaves and Putnam, 2020). Importantly, Tucker et al.'s (2019) experimental study on resumption in MSA reveal that the strongest processing effect of RPs is induced inside adjunct islands, which we assume is due to their comparatively weaker status as an island compared to RC-islands.

Considering that all experimental sentences are preceded by referential contexts, where the referent of filler phrases and RPs are mentioned, and where their salience in discourse is manipulated, we make the following predictions:

1. If RPs are discourse pronouns:
 - RPs should partially alleviate island violations
 - RPs should be sensitive to the salience of referent in discourse, i.e.:
 - They will be processed faster with a salient antecedent.
 - They will be more acceptable with a salient antecedent.
2. If RPs are syntactic variables (i.e., bound variable pronouns):
 - RPs should fully restore the acceptability of island violations.
 - RPs should not be sensitive to discourse salience.
 - RPs should facilitate processing regardless of discourse salience.
3. Individuals with lower WM capacity will be more sensitive to the presence of RPs (in terms of speed of processing, and possibly also acceptability).
4. If D-linking is not the source of the processing facilitation of *which-N* phrases on islands:
 - *Which-N* phrases should improve the acceptability and facilitate the processing of island-crossing filler-gap-dependencies irrespective of discourse salience.⁷

6.6.1 Methodology

6.6.1.1 Subjects

One hundred and sixty-three native speakers of Baha Arabic participated in this study; all participants were within the 18-24 age bracket. All were undergraduate students at the University of Baha (female section), Saudi Arabia, and received course credit for participating. The experiment was performed in a quiet room under the researcher's

⁷ Hofmeister et al. (2007) argued that, as *wh*-fillers are not anaphoric, their activation level in memory is positively sensitive to the semantic and structural complexity of the *wh*-filler, but not sensitive to the salience of their potential referents in discourse.

supervision. A language background questionnaire was completed by all participants; this was done to ensure that Baha Arabic was their dominant language, and that they had also been born in Baha or had been living in Baha for the last five years. (This is the same questionnaire as the one used in the previous studies; a translated version of this questionnaire is available in the Appendix). No participant had to be excluded on the basis of the information provided in the language background questionnaire.

6.6.1.2 Material

Experimental items were all wh-questions in which the wh-dependency crosses an adjunct island boundary. They were manipulated by crossing filler type (*ʔayy-še* ('which-N') vs. *ʔayš* ('what')) and type of dependency tail (gap vs. RP). As in our previous studies, all wh-questions in this experiment were object chains. To control for the effect of animacy, the referent of filler phrases was always inanimate. An adjunct phrase was added after the foot of the chain as a spillover region. These adjunct phrases at spillover regions were controlled for length and structure. More specifically, all three-word-long adjunct phrases were prepositional phrases (a preposition + DP).

To test the discourse properties of RPs, the experimental sentences, resulting from the factorial design described above, were preceded by context sentences that manipulated the salience of the antecedent (salient vs. non-salient). Each linguistic context consisted of three sentences. The first sentence always mentioned the critical referent as a full NP, including a modifier. The second and third sentences varied across salience conditions: in the salient condition, the critical referent is mentioned as a full NP in the second sentence and as a pronominal element in the third sentence; in the non-salient condition, the critical referent is not mentioned in either the second or in the third sentences.

The design was therefore 2 (complex vs. simple wh-phrase) x 2 (RP vs. gap) x 2 (salient vs. non-salient). The eight conditions are illustrated below.

25. A. Salient context:

Amal was poisoned after eating **the grilled prawns dish** in the lounge buffet. **The dish** was spicy, and the prawns were perfectly cooked. But unfortunately, **it** seemed to be poisoned.

Experimental sentences:

a. (What: Gap/RP)

ʔayš	amal	tsammam-at	baʕd-ma	ʔkal-at-/uh
what	Amal	poisoned-3SF	after	ate-3SF -/=it
spill over fii		bufayh	al-estrahah?	
spill over in		buffet	the-lounge	

What was Amal poisoned after she ate _/it in the lounge buffet?

b. (Which-N: Gap/RP)

ʔayy	ṭabaq	amal	tsammam-at	baʕd-ma
Which dish		Amal	poisoned-3SF	after
kal-at-/uh	spill over fii		bufayh	al-estrahah?
ate-3SF -/=it	spill over in		buffet	the-lounge?

Which dish was Amal poisoned after she ate _/it in the lounge buffet?

B. Non-Salient context:

Amal was poisoned after eating **the grilled prawns dish** in the lounge buffet. More than one person was poisoned yesterday. The council closed the restaurant that was responsible for the buffet, and fined its owner.

Experimental sentences:

a. (What: Gap/RP)

ʔayš	amal	tsammam-at	baʕd-ma	ʔkal-at-/uh
what	Amal	poisoned-3SF	after	ate-3SF -/=it
spill over fii		bufayh	al-estrahah?	
spill over in		buffet	the-lounge?	

What was Amal poisoned after she ate _/it in the lounge buffet?

b. (Which-N: Gap/RP)

ʔayy	ṭabaq	amal	tsammam-at	baʕd-ma
Which dish		Amal	poisoned-3SF	after
ʔkal-at-/uh	spill over fii		bufayh	al-estrahah?

ate-3SF -/=it spill over in buffet the-lounge?

Which dish was Amal poisoned after she ate _/it in the lounge buffet?

This self-paced reading experiment involves two sub-experiments, depending on the type of the follow-up question. In the first sub-experiment, participants were asked to rate the acceptability of the experimental sentences; in the second sub-experiment, participants were asked to perform a comprehension question task.

The distractors were all structurally complex, in that they all were bi-clausal wh-questions, but none included a crossed island boundary. The ungrammatical distractors contained either an agreement mismatch or a violation of a lexical constraint (e.g. adding a superfluous argument or violating theta-role requirements). The context sentences preceding the distractors adopt the following format:

1. A sentence that does not mention the critical referent
2. A sentence that introduces the critical referent as a DP
3. A sentence that repeats the referent as a pronoun

Some distractors include contexts in which the critical referent is introduced in the last sentence and not subsequently mentioned. Grammatical and ungrammatical filler sentences are exemplified in (26) below:

26. a. **Grammatical distractor**

- Context:

Huda likes watching TV in the afternoon. Today there was an advert for **a new hair product** that seemed to be excellent. But the presenter did not explain how to order **it**.

- Test item:

ʔayš	al-muntaj	illi	aʕjab	Muna
what	the-product that	attract.3M		Muna
baʕdma	ʕafat	al-ʔeʕlan?		
after	saw-3F		the-advertisement?	

What is the product that attracted Muna after seeing the advertisement?

b. **Ungrammatical distractor (agreement mismatch)**

- Context:

Sumaya was sick last week. She could not overcome laziness and tiredness. She went to the doctor, and he prescribed **vitamin D pills** for her.

- Test item:

ʔayš al-ʕelaj illi ad-doctorah waʕafuh

What the-medicine that the-doctor-F prescribed(M)-it

I-sumayyah?

for-Sumayyah?

*What is the medicine that the doctor (F) prescribed (M) for Sumaya?

c. **Ungrammatical distractor (violation of lexical constraint)**

- Context:

The school principal has an important meeting with her students. She wants to tell them that the registration **in a critical thinking workshop** has opened. She wants to encourage them to register for **it**.

- Test item:

ʔayy waršat ʕamal tabġa al-mudeerah

which workshop want-3F the-principal(F)

taʕreḍ maʕa ṭalbat-ha ʕan?

deliver-3F with students-her about?

*Which workshop does the principal want to deliver with her students about?

Twelve sets of lexically matched stimuli were created across the eight conditions included in the current study (with each condition lexicalised 12 times); there were eight variants of each lexicalisation in total. Experimental sentences were distributed into four counterbalanced lists using a Latin square design. Each list contained three tokens of each condition (i.e. 24 experimental sentences). Lexicalisations appeared twice per list (in a salient vs. non-salient context), separated by at least 20 sentences. Twenty-four filler items were added to each list. The full set of stimuli is presented in the Appendix.

Distributing items across lists ensures that all variables are manipulated within participants; that is, two items of each lexicalisation set appear in each list, once with salient context and once with non-salient context. The variables of Filler and Tail are

further manipulated in these two items: one item occurs with 'what' and the other item occurs with '*which-N*'; moreover, one item occurs with RP and the other with a gap.

The items within each list (48 items: 24 experimental sentences, 24 distractors) were presented in a counter-balanced order across participants so that fatigue effects would not induce a confound. Twenty participants were randomly assigned to each list: ten read the list in one order and ten read the same list but in the reverse order (i.e. 20 participants read each list).

The same lists were used in the two self-paced reading sub-experiments. These two sub-experiments are thus identical, but differ in terms of the type of the response to the follow-up task (acceptability judgment vs comprehension questions). Eighty-three participants performed the acceptability judgment sub-experiment, while 80 participants completed the comprehension question sub-experiment.

6.6.1.3 Procedures

In order to satisfy ethical research standards, full ethical approval was obtained prior to commencing the experiment (reference number: FAHC 18-049). In addition, participants were given and asked to read an information sheet, and to provide their informed consent, before taking part in the study.

Participants' WM capacity was measured to allow the inclusion of this factor as a covariate in the analyses. The Corsi block-tapping task, which was used in the previous two experiments, was also used in the present study (Stoet, 2010; Stoet, 2017). In this task, nine blocks appear on the computer screen. Sequences of blocks are then marked (lightened), beginning with a sequence of two blocks. Participants were asked to tap the 'marked' blocks in the same sequence they observed. The sequences of marked blocks increase in length as participants proceed further in the experiment. The largest correct number of marked blocks that the participant remembers is recorded as the participant's WM capacity score.

Following the completion of the individual cognitive measures, participants completed the self-paced reading experiment. The experiment was designed using the OpenSesame software (Mathôt et al., 2012). First, instructions were provided, after which four example sentences were displayed in a training session to familiarise participants with the task. In each trial, context sentences were displayed in the middle of the screen. All were presented at once (i.e. all three sentences were presented on the same screen at the same time). After reading the context sentences, participants

pressed the spacebar to proceed to the experimental sentence, which was a *wh*-question about an inanimate referent mentioned in the context sentences. The *wh*-question was presented word-by-word: i.e., one word at a time was displayed in the middle of the screen. A keypress (on the spacebar) revealed the next word in the sentence and hid the previous word.

In the comprehension question version of the experiment, participants were asked to answer the question by choosing the correct answer from two possible choices. To prevent participants from developing a response strategy by focusing on the first sentence only, the key information was provided in the second or third sentence in the distractors.

In the acceptability judgment version, participants were asked to rate the test item using a seven-point rating scale ranging from 1 ('very bad') to 7 ('very good'). The scale was displayed on a screen that appeared after participants read the test item. Participants were told to depend on their first reaction, without trying to analyse the sentence, and that there were no prescriptively 'correct' answers.

6.6.2 Results of the acceptability judgment sub-experiment

6.6.2.1 Descriptive data

We begin with an overview of the descriptive results of the acceptability judgment task. As noted above, the factors we manipulated in the present study are as follows: Salience (Salient vs Non-Salient), Filler (*What* vs *Which-N*) and Tail (RP, Gap). In this research, we are mainly interested in investigating the interaction of RPs with salience across the two dependency structures. We hypothesised that if RPs are interpreted as discourse pronouns, they will be sensitive to the salience of the antecedent, and that they will be marginally accepted regardless of filler phrase type. Moreover, if RPs are interpreted as bound variables, they should not be sensitive to salience, and should yield high levels of acceptability (as the island violation is void).

We will further investigate the extent to which *which-N* fillers ameliorate islands. If an effect is observed, as predicted by processing-based theories, we will address the nature of this processing advantage: specifically, we hypothesised that if D-linking is not the determining factor explaining why *which-N* fillers improve the acceptability of island-violating dependencies, *which-N* and *what*-questions should behave differently in gap-conditions.

Figure 6-1 plots the distribution of participants' judgments of the critical sentences, across conditions.

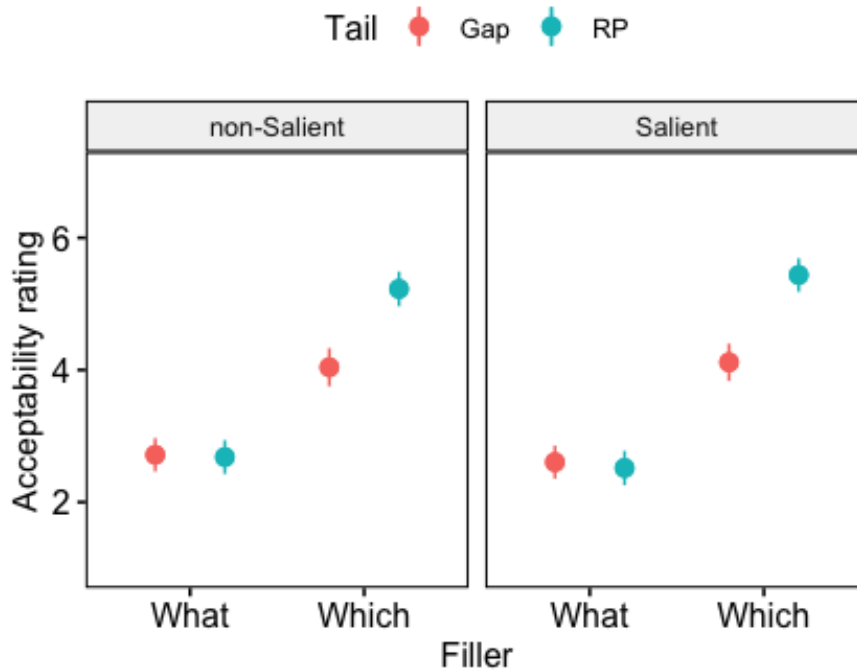


Figure 6-1: Acceptability ratings showing mean and 95% confidence intervals across structures with different fillers (What/Which), and salience (Salient vs non-Salient), with and without RP.

Acceptability scores range from 1 ('completely unacceptable') to 7 ('completely acceptable').

From visual inspection, it would appear that salience has little impact on participants' judgment of acceptability. However, there seems to be a strong two-way interaction between type of dependency filler and type of dependency tail.

Table 6-1 presents a comparison between the mean acceptability of the different conditions included in this experiment and the mean acceptability of (grammatical vs ungrammatical) filler sentences.

Table 6-1 Comparison between mean acceptability of core experimental items with fillers.

Filler phrase	Tail	Response.mean	Response.sd
What	Gap	2.66	1.98

What	RP	2.59	2.04
Which	Gap	4.07	2.25
Which	RP	5.33	2.04
Filler sentences	Response.mean	Response.sd	
Grammatical	6.17	1.73	
Ungrammatical	2.51	1.76	

6.6.2.2 Statistical analysis

Participants' acceptability judgements were recorded on a seven-point scale. To force participants to make use of the full range of the rating scale, distractor sentences included both fully acceptable and fully ungrammatical items, as stated above. However, there are still likely variations among participants that can only be partially controlled by statistical models. In particular, the distance between levels of response scale might not be perceived as identical across participants (i.e., we are dealing with an ordinal variable). To take this into account, we fitted the analysis using linear mixed effects ordinal regression models, with a cumulative link function (to permit the use of random effects). This was implemented in R (version 3.5.0) using the ordinal package (version 2019.3-9) (Christensen, 2015, R Core Team 2017). The assumption of this type of model is that the individual points on the scale are ordered, but equidistance between the points on the scale is not assumed.

6.6.2.2.1 Modelling Procedures

The models in this analysis are fitted in a bottom-up fashion, starting from random effects only (as a null hypothesis model) and incrementally adding fixed effects. The first model was constructed with Subject, Lexicalisation, and Trial Rank as random factors. The basic model was extended to include one additional variable of interest at each iteration, using likelihood ratio comparisons to ascertain whether the added variable significantly improved model fit (and discarding it otherwise). In short, fixed effects are added one by one and retained only if they improve the model's fit. Prior

to the statistical analysis, WM variable was scaled and centred to facilitate model interpretation.

Likelihood ratio tests comparing nested models revealed that the optimal model was the one that included main effects and a two-way interaction of Filler and Tail, along with random intercepts for Lexicalisation, Subject, and Trial Rank, as well as random slopes for Subject by Tail, Subject by Filler and Lexicalisation by Filler. Saliency did not improve the model fit, either as a main effect (Chi-sq= 0.9459) or in interaction with Tail (Chi-sq = 0.9835), Filler (Chi-sq = 0.342) or in a three-way interaction with these variables (Chi-sq = 0.6439). Similarly, WM did not improve the model fit, either as a main effect (Chi-sq = 0.9459) or in interaction with Tail (Chi-sq = 0.9835), Filler (Chi-sq = 0.342) or in a three-way interaction with these variables (Chi-sq = 0.6439). A summary of the optimal model is presented in Table 6-2. The summary of random effects coefficients is presented in Table 6-3, while threshold coefficients are summarised in Table 6-4.

Table 6-2 Summary of the optimal Ordinal Cumulative Link Mixed Model (clmm) of the Acceptability ratings

The formula of the model is as follows: `clmm(as.factor(response) ~ (Filler * Tail) + (1|Subject) + (1|Subject:Filler) + (1|Subject:Tail) + (1|Lexicalisation) + (1|Lexicalisation:Tail) + (1|Lexicalisation:Filler) + (1| Trial) data=datAJ)`.
Reference levels: What, Gap.

Coefficients	Estimate	Std.Er	z value	p-value
Tail				
RP	-0.099	0.125	-0.75	0.427
Filler				
Which-N	1.497	0.238	6.283	<0.001
Filler* Tail				
Which: RP	1.478	0.177	8.368	<0.001

Table 6-3 Coefficients of random effects

Groups	Variance	Std.Dev.

Subject	0.483	0.695
Lexicalisation	0.222	0.471
Subject:Filler	0.657	0.811
Lexicalisation:Filler	0.146	0.382
Trial	0.106	0.325

Table 6-4 Threshold coefficients

	Estimate	Std.Error	z value
1 2	-0.132	0.236	-0.558
2 3	0.536	0.236	2.264
3 4	1.143	0.238	4.803
4 5	1.685	0.240	7.015
5 6	2.347	0.244	9.634
6 7	3.196	0.249	12.82

For ease of interpretation, the coefficients of the model are plotted in Figure 6-2. This figure visualises the probability of each rating across conditions rather than the mean rating. Each vertical line corresponds to a combination of factor levels arising from the interaction between Filler and Tail (as determined by the optimal model). The points at which the that vertical line crosses the coloured lines indicate the probability of each rating (on the y-axis) for that specific combination of factor levels.

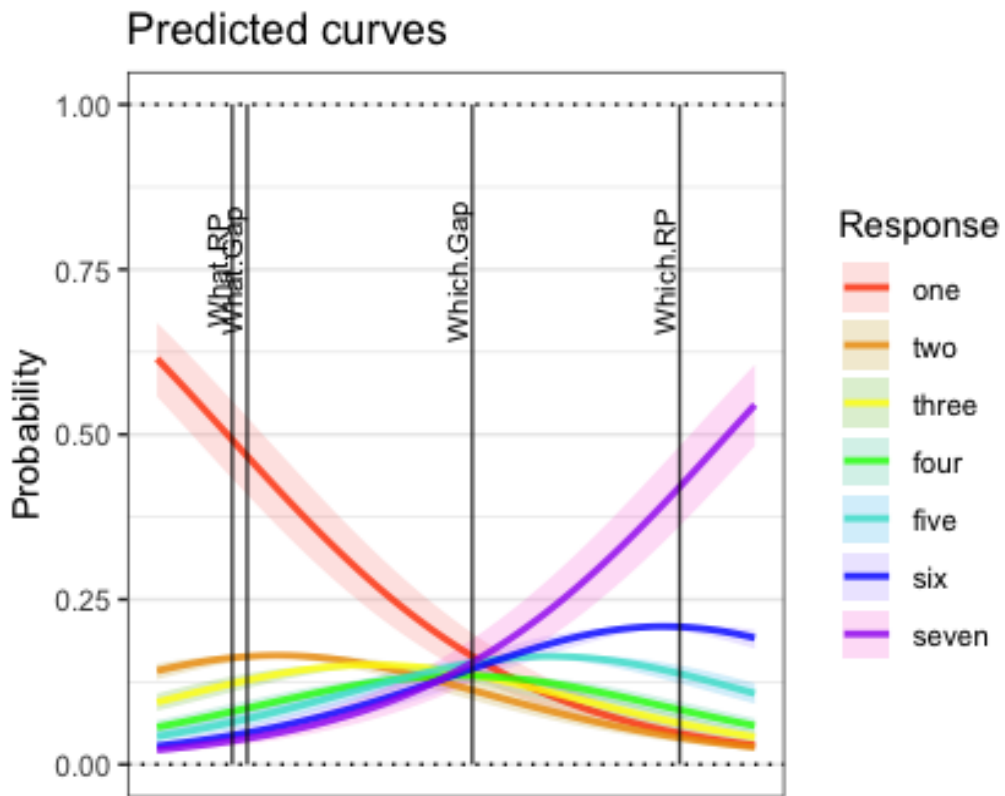


Figure 6-2: Results of the Ordinal Cumulative Link model showing the probability of each of the seven ratings in each condition (represented with vertical lines), along with Confidence Intervals (shaded).

The red line (1) represents the lowest acceptability rating, while the purple line (7) represents the highest level of acceptability.

Figure 6-2 revealed that the pattern of acceptance rate differs across conditions as follows:

- The *Which*-RP condition yielded high acceptance ratings, where the probability of acceptance (i.e. a score of 5 or more) was around 70% (adding up the probabilities for the top three ratings).
- The *Which*-Gap condition yielded mixed ratings (i.e., uncertainty), with similar levels of probability for any of the seven possible scores.
- The *What*-Gap and *What*-RP conditions yielded high levels of rejection where the probability of acceptance (i.e. a score of 3 or less) was around 70% (adding up the probabilities for the lower three ratings).

The pattern of acceptance rates across the experimental results revealed a significant two-way interaction effect between Filler and Tail. In what-questions, RPs are rejected

as often as gaps, i.e. there is no significant difference between gaps and RPs in what-questions (Estimate: -0.099, z: -0.75, p: 0.427); in other words, RPs do not significantly ameliorate island effects. In *which*-questions, on the other hand, RPs significantly improved acceptability (Estimate: 1.47, z: 8.368, p: <0.001).

Furthermore, *which*-questions with gaps are preferred to what-questions with gaps (Estimate: 1.49, z: 6.28, p: <0.001), suggesting that *which*-N fillers have an independent processing effect on islands.

However, as seen in Figure 6-2, *which*-questions with gaps received inconsistent rating patterns, being accepted and rejected by participants to an equal extent. Closer investigation of participants' ratings for this condition revealed a high amount of individual variation. Figure 6-3 also shows strong variability within participants: the same participants rated sentences in this condition as acceptable, unacceptable, and somewhere in the middle. Furthermore, it would appear that this variability is not due to the lexicalisation effect; as observed in Figure 6-4, experimental sentences of different lexicalisation sets were not associated with a consistent rating pattern. Moreover, Figure 6-5 shows that variability within participants is not conditioned by lexicalisation.

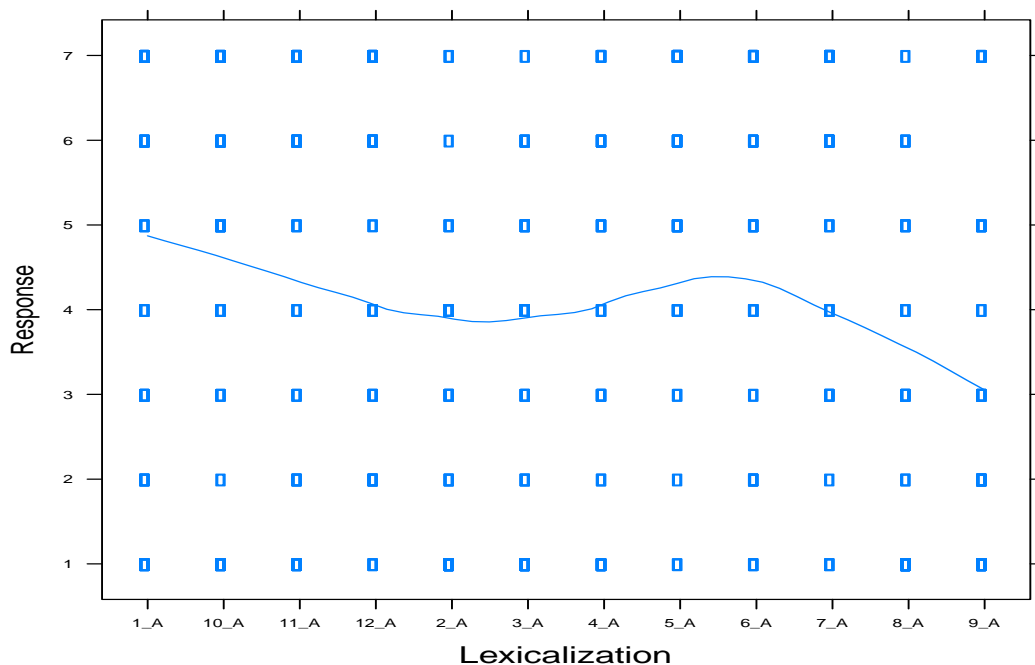


Figure 6-3: Variability in the ratings received by each lexicalisation in *which*-questions with gaps.

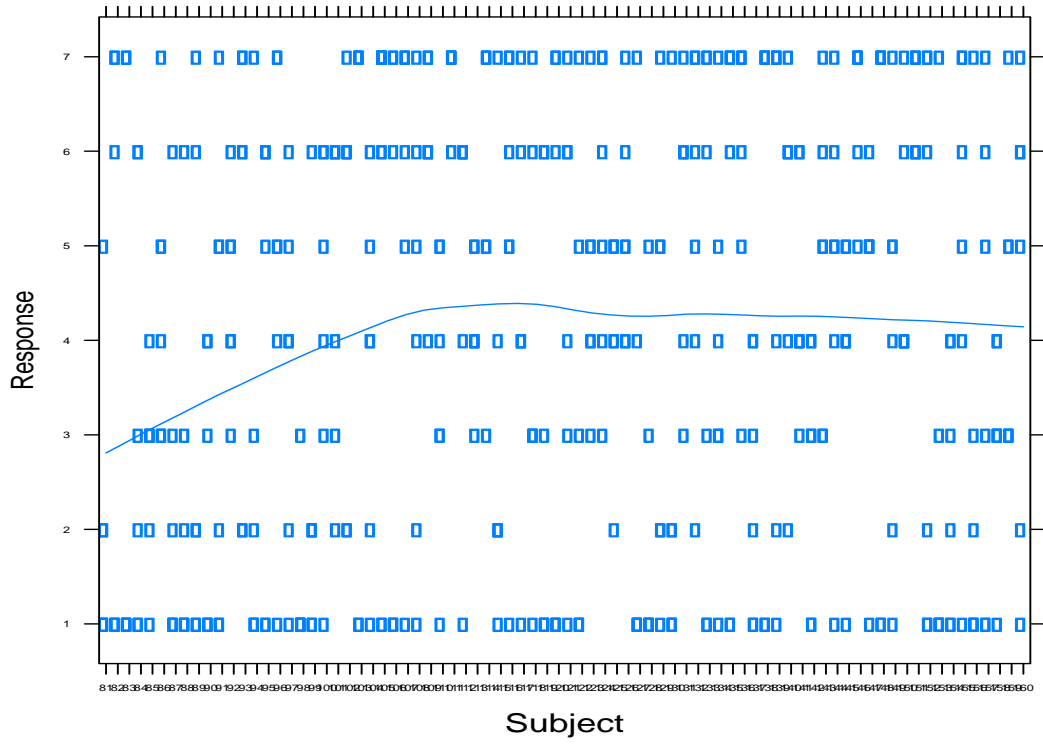


Figure 6-4 Variability in the rating of *which*-questions with gaps by subjects.

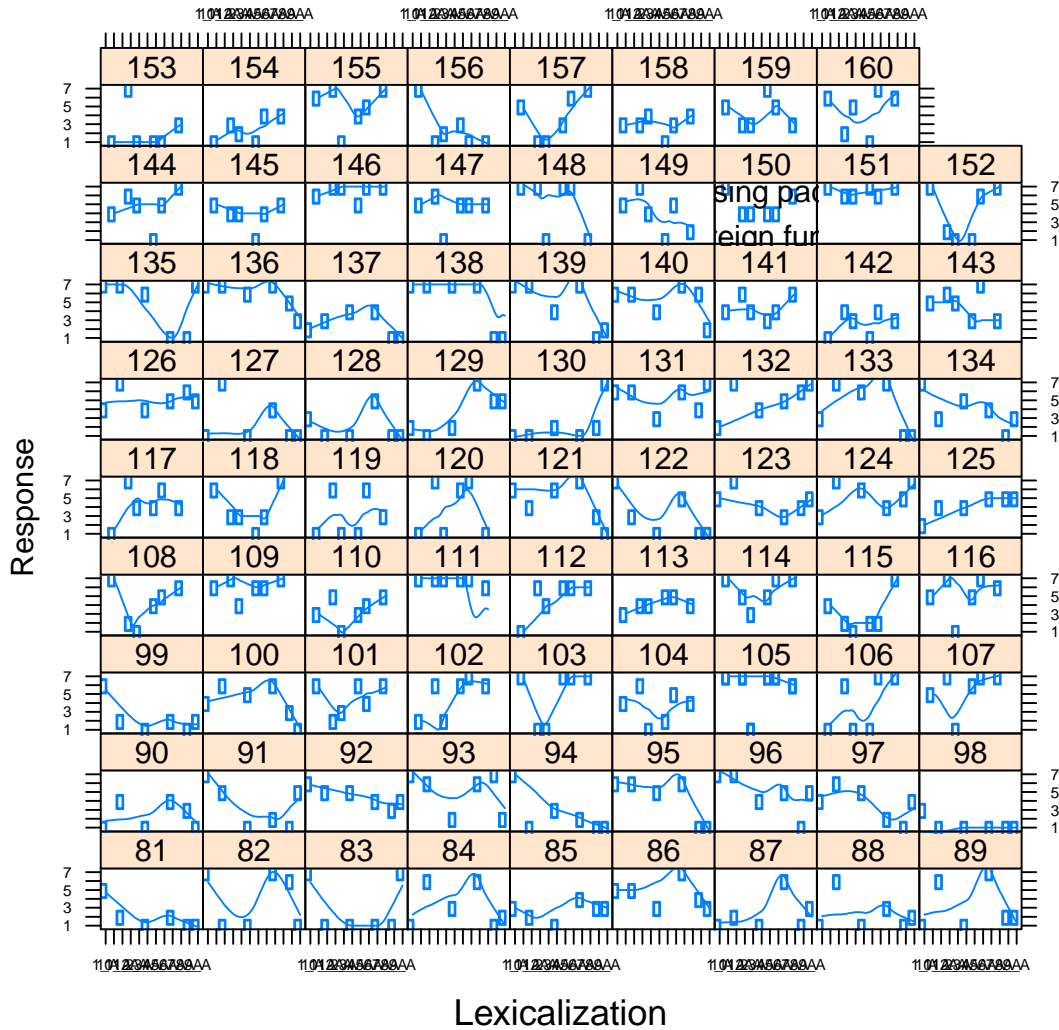


Figure 6-5 Variability in the rating of *which*-questions with gaps by items across subjects.

6.6.2.3 RT data for the acceptability judgment sub-experiment

6.6.2.3.1 Descriptive data

In this study, we evaluate reading times at the spillover region following the RP or gap, which comprises an adjunct phrase consisting of three words (bolded in 27):

- | | | | |
|-----|-----------------|---------------|----------------------|
| 27. | ʔayš/?ayy ʔabaq | tsammamat | Amal baʕd-ma |
| | what/which dish | poisoned.PASS | Amal after |
| | ʔkalat-/uh | fi | bofyyat |
| | ate-/it | in | buffet |
| | | | al-ʔestrahah? |
| | | | the-lounge? |

What was poisoned Amal after she ate _/it in the lounge buffet?

Notably, we did not depend on RT at the integration position (subcategorising verb); this is because the presence of RPs, which are cliticised to the verb, will independently lead to an increase in RTs of verbs with RPs in comparison to verbs with gaps. Mean raw reading times for the average of the three words in the spillover region are plotted in Figure 6-6.



Figure 6-6 Average of raw RT of the three words in the spillover region, showing mean and 95% confidence intervals across questions with different types of fillers (What vs Which) with and without RP.

For clarity sake, Table 6-5 shows the mean raw RTs at integration site and the mean average raw RTs at spillover region across conditions. Note that there is a pronoun linked to the subcategorization verb in RP-dependencies as opposed to gap-dependencies.

Table 6-5 The mean raw RTs at integration site and the mean average raw RTs at spillover region across conditions.

Variable 1	Variable 2	RT at the verb region	Average RT at spill over region

Tail	Filler	RT.mean	RT.sd	RT.mean	RT.sd
Gap	What	499.50	256.18	551.50	414.67
RP	What	527.62	333.39	559.33	403.27
Gap	Which	526.03	355.95	611.80	439.15
RP	Which	562.92	456.97	544.64	373.17

6.6.2.3.2 Statistical analysis

Reading times were analysed via Generalised Additive Models (Wood, 2011), using the *mgcv* package in R studio (Version 1.1.419). This model was selected because it takes the non-linear nature of reading time data into consideration. Prior to statistical analysis, raw RT scores with values greater than 5000 ms and smaller than 150 ms were removed from the data. This was done because values less than 150 ms likely indicate that participants did not read the word (either accidentally or on purpose); moreover, values above 5000 indicate that participants might have been distracted, experiencing technical difficulties, or over-thinking. Forty data points were removed as a consequence of this procedure (20 points in gap-conditions and 17 points in RP-conditions in *which*-questions, and 27 points in gap-conditions and 23 points in RP-conditions in *what*-questions).

Prior to statistical analysis, raw reading times at the regions of interest were box-cox-transformed to normalise the data. The distribution of the box-cox-transformed RT data is presented in Figure 6-7.

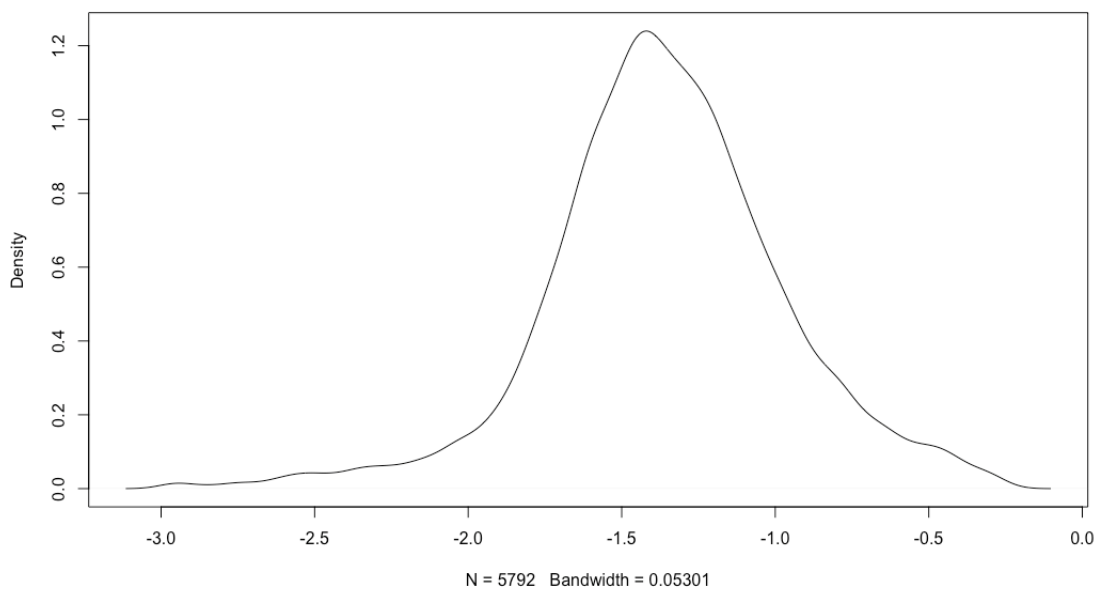


Figure 6-7 Distribution of box-cox-transformed RT data.

The models in this analysis are fitted in a bottom-up fashion, starting from random effects only (as the null hypothesis model) and incrementally adding fixed effects. The first model was constructed with Subject, Lexicalisation and Trial Rank as random factors. The basic model was extended to include one additional variable of interest at each iteration, using likelihood ratio comparisons to ascertain whether the added variable significantly improved model fit (and discarding it otherwise). Prior to the statistical analysis, the WM variable was scaled and centred to facilitate model interpretation.

Likelihood ratio comparisons between nested models revealed that the optimal model formula was the same as that for the acceptability judgement data: the optimal model includes main effects and two-way interaction of fixed effects (Filler, Tail), as well as random intercepts for Lexicalisation, Subject and Trial, and random slopes for Subject by Tail, Subject by Filler, Lexicalisation by Tail and Lexicalisation by Filler. Salience did not improve the model fit, either as a main effect or in interaction with Tail, Filler or in a three-way interaction, as indicated by compareML. Similarly, centered.WM did not improve the model fit.

The summary of the optimal model is shown in Table 6-6.

Table 6-6 Summary of the optimal Generalized Additive Model (GAM) of the RT data in the Acceptability Judgment sub-experiment.

The formula of the model is: $RT.t \sim (Filler * Tail) + (1|Subject) + (1|Subject:Filler) + (1|Subject:Tail) + (1|Lexicalization) + (1|Lexicalisation:Tail) +$

(1|Lexicalisation:Filler) + (1| Trial) data=datRT). Reference levels: What/Gap.

Coefficients	Estimate	Std.Er	t value	p-value
Intercept	-1.663	0.034	-48.239	<0.001
Filler				
What	--	--	--	--
Which	0.092	0.018	4.917	<0.001
Tail				
Gap	--	--	--	--
RP	0.033	0.016	2.119	0.034
Filler * Tail				
Which:RP	-0.073	0.016	-4.442	<0.001

Similar to the acceptability judgment data, the RT data shows that salience of the antecedent in the preceding context does not affect RP processing in either in what-questions or *which*-questions. Results revealed a two-way interaction between Filler and Tail, with *which*-questions with gaps being slower than what-questions with gaps at the spillover region (Estimate: 0.092, t: 4.917, p: <0.001). RPs affect the processing of both types of questions in different ways; specifically, results revealed that RPs significantly slow down RTs in what-questions (Estimate: 0.033, t: 2.11, p: <0.03), but speed up RTs in *which*-questions (Estimate: -0.073, t: -4.44, p: <0.001).

6.6.3 RT data in comprehension question data

In this self-paced reading sub-experiment, experimental items, all of which are wh-questions, were followed by a multiple-choice task. Participants were asked to choose the correct answer to the experimental item wh-question from among two choices.

6.6.3.1 Descriptive data

As in the previous studies, we will not focus on the subcategorizing verb as our region of interest since comparing RTs in gap dependencies and RP dependencies at this region will necessarily be longer if there is a pronoun as opposed to no lexical material. Instead, we evaluate reading times at the spillover region following the RP or gap, which comprises an adjunct phrase consisting of three words. However, for clarity sake, Table 6-7 shows the mean raw RTs at integration site and the mean average raw RTs at spillover region across conditions.

Table 6-7 The mean raw RTs at integration site and the mean average raw RTs at spillover region across conditions.

Variable 1	Variable 2	RT at the verb region		Average RT at spill over region	
Tail	Filler	RT.mean	RT.sd	RT.mean	RT.sd
Gap	What	340.10	202.97	372.58	167.89
RP	What	340.35	151.11	376.97	183.32
Gap	Which	341.76	151.96	377.09	190.65
RP	Which	333.37	110.51	369.75	127.91

6.6.3.2 Statistical Analysis

Prior to statistical analysis, raw RT scores with a value greater than 5000 ms and smaller than 150 ms were removed from the data. This was done because values below 150 ms likely indicate that participants did not read the word (either accidentally or on purpose); moreover, values above 5000 indicate that participants might be distracted, meaning that conscious processing has not taken place. Forty-seven data points were removed as a consequence of this procedure: seven points in gap-conditions and 10 points in RP-conditions in *which*-questions, and 17 points in gap-conditions and 13 points in RP-conditions in *what*-questions. No participant had been

excluded from the statistical analysis based on the mean question-answer accuracy. Accuracy rate was always higher than 80% for both experimental sentences and filler sentences across all subjects.

In addition, prior to statistical analysis, raw reading times at the regions of interest were box-cox-transformed to normalise the data. The distribution of the box-cox RT data is presented in Figure 6-8.

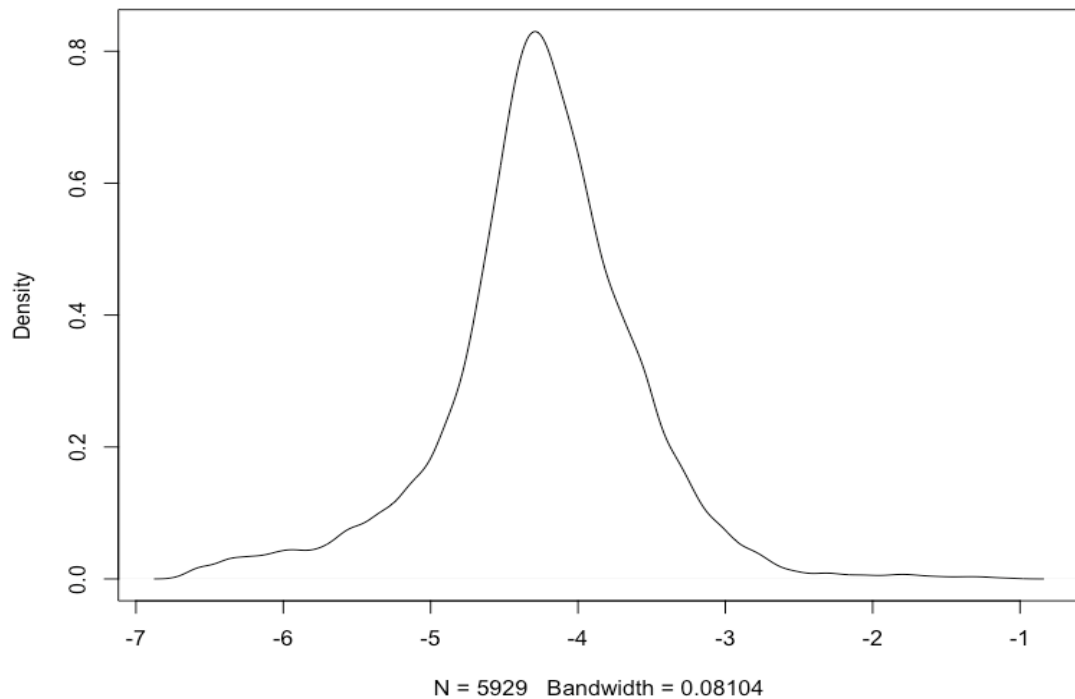


Figure 6-8 Distribution of box-cox RT data.

Data were analysed with Generalised Additive Models (Wood, 2011), using the *mgcv* package (Version 1.8.25). in R Studio (Version 3.5.0). This model was selected because it takes the non-linearity of reading time data into consideration. The models in this analysis are fitted in a bottom-up fashion, starting from random effects only (as a null hypothesis model) and incrementally adding fixed effects. Likelihood ratio comparisons between nested models, conducted via the *fREML* method, revealed that adding Tail, Filler, Salience and centred WM to the null hypothesis model do not improve model fit, either as main effects or as interactions. The formula of the model with the best fit is as follows: $RT.t \sim (1|Subject) + (1|Lexicalisation) + (1| Trial)$ data=datRT). The compareML results are reported in Table 6-8.

Table 6-8 Results of Nested Model comparisons as revealed by compareML method.

Model fit	Preferred model compareML (outcomes)
1. RT.t ~ (1 Subject) + (1 Lexicalisation) + (1 Trial) data=datRT)	The null hypothesis model
2. RT.t ~ (1 Subject) + (1 Lexicalisation) + (1 Trial) + Filler, data=datRT)	Model 1
3. RT.t ~ (1 Subject) + (1 Lexicalisation) + (1 Trial) + Tail, data=datRT)	Model 1
4. RT.t ~ (1 Subject) + (1 Lexicalisation) + (1 Trial) + (Tail:Filler), data=datRT)	Model 1
5. RT.t ~ (1 Subject) + (1 Lexicalisation) + (1 Trial) + Saliency, data=datRT)	Model 1
6. RT.t ~ (1 Subject) + (1 Lexicalisation) + (1 Trial) + (Saliency:Filler), data=datRT)	Model 1
7. RT.t ~ (1 Subject) + (1 Lexicalisation) + (1 Trial) + (Saliency:Tail), data=datRT)	Model 1
8. RT.t ~ (1 Subject) + (1 Lexicalisation) + (1 Trial) + (Saliency:Tail:Filler), data=datRT)	Model 1
9. RT.t ~ (1 Subject) + (1 Lexicalisation) + (1 Trial) + centered.WM, data=datRT)	Model 1
10. RT.t ~ (1 Subject) + (1 Lexicalisation) + (1 Trial) + (centered.WM:Filler), data=datRT)	Model 1
11. RT.t ~ (1 Subject) + (1 Lexicalisation) + (1 Trial) + (centered.WM:Tail), data=datRT)	Model 1

12. RT.t ~ (1 Subject) + (1 Lexicalisation) + (1 Trial) + (centered.WM:Saliency), data=datRT)	Model 1
13. RT.t ~ (1 Subject) + (1 Lexicalisation) + (1 Trial) + (centered.WM:Filler:Tail), data=datRT)	Model 1
14. RT.t ~ (1 Subject) + (1 Lexicalisation) + (1 Trial) + (centered.WM:Filler:Tail:Saliency), data=datRT)	Model 1

In summary, the present results revealed that, in contrast to the RT data obtained from the acceptability judgment task, neither Tail nor Filler type has an impact on processing island-violating dependencies. The difference between the results of this experiment and the acceptability judgment experiment suggests that the type of the follow-up task affects the processing of experimental sentences in RT data. See Figure 6-9, which plots mean raw reading times at the spillover region across conditions in the two different tasks.

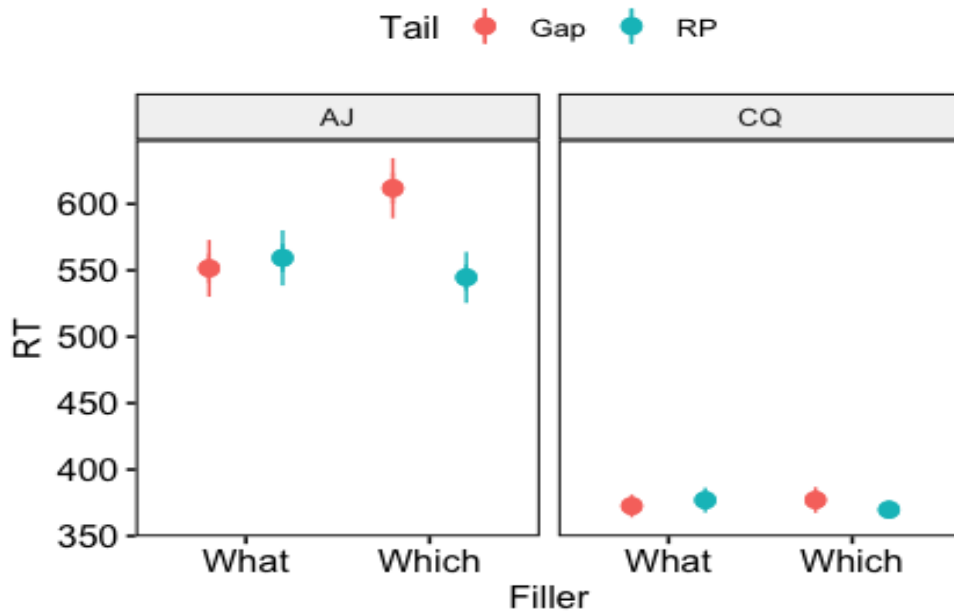


Figure 6-9 Plot with the average raw RT at the three-word spillover region, showing mean and 95% confidence intervals across questions with different types of fillers (What vs Which) with and without RP, in the two different tasks.

6.7 Discussion

This experiment was designed to ascertain the extent to which RPs in Baha Arabic can alleviate island violations in inherently and contextually D-linked wh-questions when presented in a felicitous context. Moreover, we also wanted to determine the extent to which *ʔayya-še* ('which-N') fillers can ameliorate island effects in the absence of RPs. Finally, we manipulated the type of the task accompanying the self-paced reading task (i.e., a judgement or comprehension question) to assess whether task type might have had a confounding effect in the previous studies.

We carried out an online self-paced reading task manipulating the type of element at the integration site (RP vs Gap) in *ʔayš* ('what') questions vs *ʔayya-še* ('which-N') questions. A carefully controlled context provided an explicit referent for the wh-filler phrase, manipulating the discourse salience of this referent.

The syntactic and processing theories of D-linking both make the similar prediction that RPs should be preferred to gaps, irrespective of the source of D-linking (i.e., contextual or inherent) in island-violating wh-questions. However, these theories make different predictions in regard to (i) the magnitude of the amelioration effect of RPs across the two types of D-linked wh-questions and (ii) the sensitivity of resumption to contextual discourse salience. The syntactic account considers RPs to be bound pronominal variables in syntactic-binding dependencies; thus, it predicts that the presence of RPs will restore island-violating dependencies to full acceptability. By contrast, the processing account argues that RPs are interpreted as discourse pronouns capable of anaphorically resolving ungrammatical syntactic-movement dependencies; thus, it predicts that RPs will only partially improve the acceptability of island-violating wh-questions by alleviating the processing demands of island structures and allowing an interpretation of the dependency.⁸ Consequently,

⁸ Kluender and Kutas (1993) and Hofmeister and Sag (2010) proposed a purely processing-based theory for island effects. However, as we explained in Chapter 3, the empirical facts and experimental findings both support the view that island effects should be attributed to both syntactic and non-syntactic factors. Indeed, the results of the first experimental study strongly supported the syntactic distinction between movement and binding dependencies, with an indication that processing factors also contribute to the picture. Thus, the view of islands we adopt here is one that incorporates both syntactic and cognitive factors (Goodall, 2014; Keshev and Meltzer-Asscher, 2019).

the processing account predicts that RPs will be sensitive to the salience of the referent in discourse, while the syntactic accounts make no such prediction.

The experimental results revealed that RPs only improve acceptability and facilitate processing of island-violating *wh*-questions with *ʔayya-še* ('which-N') fillers, but not in *what*-questions. Neither the processing nor the syntactic account of D-linking predicts this interaction between resumption and filler type if D-linking is enabled by the context.

However, as we explained in 6.2, the possibility of RP licensing by the discourse context varies depending on the type of filler phrase in Arabic. For instance, Aoun and Choueiri (1999) argue that *šu* ('what') and *maḡaa* ('what') cannot discourse-link in Lebanese Arabic and MSA respectively, and therefore cannot license RPs. Meanwhile, Abdel Razaq (2011) claims that while *šu* ('what') cannot be contextually discourse-linked in Jordanian Arabic, while *ʔayš* ('what') can be contextually discourse-linked, and therefore licenses RPs similarly to inherently D-linked fillers. The findings of the current study suggest that, in Baha Arabic, *ʔayš* ('what') questions cannot be contextually discourse-linked, as discourse licensing of RPs appears to be restricted to dependencies with inherently D-linked fillers (e.g. *ʔayya-še* ('which-N')).

The implications of this finding are discussed below for syntactic theories (in 6.7.1) and for processing theories (in 6.7.2).

6.7.1 Syntactic theories

The assumption that D-linking interpretation is only licensed in *which*-questions takes us back to the discussion of whether these two types of questions (i.e., *ʔayš* ('what-questions') and *ʔayya-še* ('which-questions')) have the same or different underlying syntactic structures. One proposed theory is that movement chains are involved in the derivation of *what*-questions (which explains their sensitivity to islands); in *which*-questions, on the other hand, RPs are externally merged from the lexicon and are obligatorily interpreted as co-referent with the filler phrases at LF (which bypasses island constraints). This is the standard analysis adopted by Arabic syntacticians for resumption in both relative clauses and D-linked *wh*-questions (Aoun et al., 2009).

The syntactic account assumes that, similarly to relative clauses, the pronoun in *ʔayya-še* ('which-N') questions is externally merged from the lexicon and obligatorily bound by the filler phrase at LF. This syntactic account therefore makes the following predictions: (i) that gap-dependencies are ungrammatical, similarly to relative clauses

with gaps; (ii) that RP-dependencies are grammatical; and finally (iii) *ʔayš* ('what') questions are unacceptable both with or without RPs, as they involve a violation of grammatical constraints on movement. At first sight, these predictions appear to be confirmed by our results: in island-violating configurations, the presence of a RP significantly improves the acceptability of *ʔayya-še* ('which-N') questions, but not *ʔayš* ('what') questions.

However, *ʔayya-še* ('which-N') questions with gaps were found to exhibit an inconsistent pattern of acceptability, including pervasive inter-individual variation. This variability is unexpected under a binding analysis, which predicts ungrammaticality due to the violation of a subcategorisation requirement. The variation in acceptability observed in *ʔayya-še* ('which-N') questions with gaps also differs substantially from that of relative clauses with gaps in Study 1 (which were consistently rejected).⁹ To some extent, the parser appears able to derive a representation for *ʔayya-še* ('which-N') dependency despite the lack of RP. This could be due to some kind of accommodation facilitated by the filler phrase.

We conclude that the underlying syntactic structure of *ʔayya-še* ('which-N') questions is different from that of *illi*-dependencies: syntactic binding is only involved in the latter.

Which-questions with gaps are predicted to be sensitive to grammatical constraints on movement (i.e. island effects). The inconsistent pattern of acceptability of *which*-questions with gaps observed in our results could be due to the impact of non-syntactic factors on the acceptability of island-violating dependencies. Indeed, the acceptability of island-violating dependencies is conditioned by a variety of cognitive constraints in addition to syntactic constraints (Goodall, 2014; Keshev and Meltzer-Asscher, 2019; Chaves and Putnam, 2020; Pañeda et al., 2020; Perpiñán, 2020). The nature of this amelioration effect will be discussed further in 6.7.2.2. For now, it is sufficient to note that the syntactic account's contention that *illi*-structures and *which*-questions involve the same type of dependency is incompatible with our results. One important consequence of this analysis, in line with Alexopoulou's (2010) theory, is

⁹ It could be the case that the absence of plausible context in the first acceptability judgment has had an effect here. However, the contrasts we observed between relative clauses and *ʔayya-še* ('which-N') questions in online comprehension studies, as explained below, suggest that the absence of plausible context is unlikely to account for the contrast in acceptability judgment experiments.

that these two types of structures involve two different types of RPs: namely, RPs are intrusive in *ʔayya-še* ('which-N') questions, but are true RPs in relative clauses.

Further support for the different underlying syntactic structures in relative clauses vs *ʔayya-še* ('which-N') questions in BA, comes from their different behaviour in the self-paced reading experiments with comprehension questions. Only self-paced reading tasks with judgements, but not those with comprehension questions, were able to detect differences between *which*-questions with gaps vs those with RPs. However, a self-paced reading task with comprehension question was able to detect differences between relative clauses with gaps vs RPs (in Study 2). This suggests that parsers are more sensitive to the absence of grammatical RPs in relative clauses (a violation of grammatical rules) than to the presence of intrusive RPs in *which*-questions (a facilitation effect).

These observations are compatible with a 'last resort' syntactic account (Shlonsky, 1992; Aoun, 2000) of RPs in *which*-questions in Baha Arabic (in contrast with a first-merge account of RPs in relative clauses). As RPs only improve D-linked *which*-questions, but not non-D-linked what-questions, the 'last resort' insertion of RPs is only licensed with inherently D-linked filler phrases, consistent with Aoun et al. (2009).

Despite the popularity of this last resort account in syntactic theories, there are still problems in implementing this theory from a technical perspective. Aoun et al. (2001) postulated a derivational implementation for binding relation for RPs in Lebanese Arabic wh-dependencies; specifically, resumptive pronouns are introduced in the derivation whenever movement takes place inside islands (i.e. as last resort devices). Their assumed 'Bind' operation involves demerging the operator and re-merging it at the final destination site, while a pronominalisation process takes place at integration site (after all, resumptive pronouns are interpreted as syntactic variables bound by the operator at C). Importantly, the Bind operation is only resorted to when a violation of grammatical constraints exists; in cases when no such violation occurs, the Bind operation is blocked, as it is more costly than the Move operation (since it involves more derivational steps and requires a pronominalisation process).

However, the mechanisms in which the alternation between movement and binding derivations take place face different problems. For instance, it is well known that adding pronouns to the numeration while the derivation proceeds would constitute a violation of Chomsky's (1995) inclusiveness condition, which prohibits introducing new elements not previously present in the numeration during the derivation. Among other problems for such accounts is that it is unclear how the feature, for which the

filler moves initially to satisfy, will be satisfied later if a pronoun is inserted, and the dependency is resolved by binding relations. Moreover, what about the intermediate copies? According to Salzmann (2009), this account predicts resumptive pronouns in intermediate positions in island-violating dependencies where the operator moves before encountering the extraction site inside islands, as in (28).

28. * $[_{CP} C_{+wh} \quad [_{ISLAND} [_{CP} Op_i \quad [_{CP} Op_i \dots Op_i]]]]]$.
↓ res

(Salzmann, 2009, p.66)

According to Martorell (2018), the syntactic operation Merge (Chomsky, 1995) 'is strictly conceived at the competence domain as an offline computation without real-time implications' (p.2); one piece of evidence for this fact is that 'Merge generates sentences from right to left, namely in the opposite direction to that required by language processing' (Martorell, 2018, p.2). Hence, competence-based theories provide no account of how words are combined hierarchically in real-time. However, it is commonly accepted that the results of experimental studies on this subject align with the considerations of theoretical linguistics (Phillips and Wagers, 2007; Lewis and Phillips, 2014): that is, ungrammatical sentences tend to take longer to process at regions of interest when compared to grammatical sentences.

Martorell (2018) proposed the predictive structure-building mechanism to account for how syntactic information can be reflected in real-time processing. This mechanism involves two computational stages: 'a top-down stage for predicting syntactic structure and a bottom-up stage for syntactically integrating actual input' (Martorell, 2018, p.2). Martorell (2018) further assumed that the top-down syntactic prediction is limited to core syntactic information (i.e. verbs and their arguments); for instance, processing a transitive verb will trigger the expectation that an object will follow, and the absence of such an object will slow real-time processing.

Adopting the view of Martorell's (2018) predictive structure-building mechanism, the last resort syntactic account would argue that the slowdown in RTs in *ʔayya-še* ('which-N') questions with gaps reflects participants' unmet expectation of encountering an object for the verb (as the grammatical constraints on movement do not allow filler phrases to be interpreted when the integration site is located inside an island). The facilitation effect of RPs in *ʔayya-še* ('which-N') questions therefore occurs because the presence of RPs satisfies the theta-role requirement and a binding relation is established between the filler and the pronoun.

However, it is unclear under the syntactic theory why *ʔayya-še* ('which-N') questions are accepted in absence of RPs. This is contrary to the expectations established by syntactic accounts of islands and indicates that parsers are able to reactivate the filler phrases inside islands; accordingly, this suggests that non-syntactic factors, induced by filler phrase type, make a contribution here.

6.7.2 Processing theories

Under a processing theory, both resumption and *which-N* fillers are expected to facilitate the parsing of island-violating structures. We discuss each in turn below.

6.7.2.1 Resumption as a processing phenomenon

It has been argued that the discourse properties of intrusive RPs enable parsers to interpret dependencies anaphorically when the derivation of the syntactic dependency fails (Alexopoulou, 2010; Chacón, 2019). Importantly, this account assumes that both *ʔayš* ('what') questions and *ʔayya-še* ('which-N') questions involve the same type of filler-gap dependency (i.e. movement dependency). However, the D-linking properties of filler phrases and the types of element at dependency tail affect the ease with which these structures are processed.

The observation that resumption interacts with the D-linking properties of filler phrases has always been taken as evidence supporting the claim that RPs are processed just like discourse pronouns in FGDs. However, in the current study, we have attempted to dissociate the two, manipulating discourse salience independently from filler type and the presence of RPs.

If RPs, given their discourse properties, are assumed to behave like pronouns, they can thus be expected to be more acceptable in the presence of a prominent, highly accessible antecedent in the preceding discourse (Ariel, 1990; Patterson, 2013). Thus, if RPs inside islands are processed as discourse pronouns in D-linked *ʔayya-še* ('which-N') questions, their acceptability and processing will be sensitive to the salience of the referent in discourse; i.e., the strongest facilitation effect of RPs in *ʔayya-še* ('which-N') questions is expected to be observed when preceded by a contextually salient antecedent in discourse. However, this prediction was not confirmed in either the processing or in the acceptability experiments (that is, manipulating the salience of the referent in discourse neither improves acceptability nor facilitates processing of RPs in *ʔayya-še* ('which-N') questions).

However, before it can be concluded that RPs in Baha Arabic wh-questions are insensitive to discourse salience, and are therefore not interpreted as discourse pronouns, a baseline for sensitivity to discourse salience with discourse pronouns in Baha Arabic should be established; such a baseline is not available in the current study. Nevertheless, the relatively high acceptance rate of RPs in *which*-questions still suggests that the amelioration effect of RPs occurs at the grammatical level, indicating that RPs are not interpreted as discourse pronouns but rather as bound variables.

6.7.2.2 Type of filler phrase

One of our aims in this study is to determine whether or not the acceptability and processing of islands is sensitive to the type of filler phrase in Baha Arabic. As illustrated in 6.7.1, our results revealed that island-violating *ʔayya-še* ('which-N') questions with gaps received an inconsistent pattern of acceptability ratings with strong inter-individual variation, while island-violating *ʔayš* ('what') questions with gaps were consistently rejected. As a rule, the gradient and the variation in the acceptability of island-violating dependencies is commonly viewed as evidence to support the processing-based accounts of islands, which link island effects to limitations on WM resources (Kluender and Kutas, 1993b; Hofmeister and Sag, 2010) and/or to the pragmatic/semantic factors that contribute to the plausibility of the interpretation of the island-violating dependencies (Erteschik-Shir, 1973; Erteschik-Shir and Lappin, 1979; Goldberg, 2006; Chaves and Putnam, 2020).

In this section, we will discuss the observed impact of filler type on islands from the perspective of the non-syntactic theories: specifically, the D-linking account (Frazier and Clifton, 2002) and the semantic and structural complexity account (Hofmeister and Sag, 2010).

The D-linking processing theory (Frazier and Clifton, 2002) is based on the assumption that dependencies in island-violating D-linked wh-questions are resolved anaphorically, allowing parsers to arrive at a sensible interpretation of the dependency regardless of its grammatical status. Although no differences between inherently and contextually D-linked questions are expected in this regard, the results have revealed that island-violating wh-questions in the gap condition behave differently: *ʔayya-še* ('which-N') questions received higher acceptability ratings and were processed slower than *ʔayš* ('what') questions.

We proposed in 6.7 that the observation that RPs are only licensed in *ʔayya-še* ('which-N') questions suggests that *ʔayš* ('what') questions cannot be contextually D-linked, and further that the discourse-linking interpretation is restricted to dependencies with inherently D-linked fillers (e.g. *ʔayya-še* ('which-N')). Thus, from the viewpoint of the processing account of D-linking, the improvement in the acceptability of *ʔayya-še* ('which-N') questions with gaps can be attributed to their D-linking properties, which enable parsers to interpret illicit syntactic dependencies anaphorically; the increased slowdown in RTs at the spillover region in *ʔayya-še* ('which-N') questions with gaps compared to *ʔayš* ('what') questions with gaps thus reflects parsers' attempt to establish a link to a previous referent in discourse (i.e. access the discourse representation to retrieve information about the referent of the filler phrase) (Frazier and Clifton, 2002). However, our data are still not fully compatible with this view.

If the dependency in *ʔayya-še* ('which-N') questions with gaps is indeed resolved anaphorically, the salience of the referent in the discourse is expected to have an impact on the acceptability and processing of these structures (Troyer et al., 2016); at least, it is expected that the inconsistent pattern of acceptability will be reduced when the discourse referent of the filler is salient in discourse. Kluender and Kutas (1993) and Kluender (1998) view D-linking as a means of 'boosting' the activation level of the filler phrase in memory. Specifically, these authors contend that the activation level of D-linked *wh*-fillers in memory is boosted when they are further associated with higher salience and prominence in discourse, leading to successful maintenance of the filler in memory until the gap is encountered. This theory is closely related to Ariel's (1990) accessibility hypothesis, which states that salient referents in discourse have a robust memory advantage during retrieval in resolving anaphoric dependencies. From this perspective, salient (i.e. highly accessible) referents are expected to both facilitate processing and improve the acceptability of D-linked *which*-questions. However, this was not found to be the case in the present research: manipulating the salience of the referent in discourse impacted neither the acceptability nor the processing of *ʔayya-še* ('which-N') questions with gaps.

One might hypothesise that this inconsistent pattern of acceptability is due to the failure to assign an anaphoric interpretation by some speakers or in some contexts. However, it has been observed that variability is conditioned neither by participants nor by item (see Figure 6-3, Figure 6-4 and Figure 6-5), and that it is observed even

in inherently D-linked questions that are preceded by a context promoting the salience of the target referent; both of these observations are problematic for this hypothesis.

Moving on to the complexity-based processing theory, Hofmeister and Sag (2010) argued that the distinguishing property of *which*-N phrases that allows them to ameliorate island effects is their richer structural and semantic content in comparison to bare *wh*-phrases. Although semantically and structurally complex *wh*-phrases are more costly to process, this processing cost translates into an increased activation level of the filler in memory, which in turn makes it more accessible for retrieval for integration at the gap site (Kluender, 1991; Hofmeister and Sag, 2010; Goodall, 2014).

Importantly, Hofmeister et al. (2007) argued that D-linked *wh*-fillers are not anaphoric, although their interpretation may derive in part from a preceding discourse. Thus, Hofmeister et al. (2007) argued that, unlike referential NPs (whose activation level in memory is sensitive to the salience of their referent in discourse), the activation level of *wh*-fillers in memory is only sensitive to the semantic and structural complexity of the *wh*-filler. Consequently, these authors argued against the relevance of discourse-related factors to the processing of FGDs. Hofmeister and Sag (2010) provided evidence to support this hypothesis by demonstrating that complex non-referential fillers ameliorate island effects to a greater extent than simple non-referential fillers.

The insensitivity of island-violating dependencies to the salience of the referent in discourse, as well as their insensitivity to contextually D-linked fillers (if they discourse link), can thus be taken as evidence in support of the complexity-based processing theory over the D-linking processing theory.

However, the fact that RTs slow down at the spillover region in *ʔayya-še* ('*which*-N') questions is not consistent with the assumptions of this theory; that is, this hypothesis is based on the observation that *which*-N fillers in English are associated with an increased ease of processing at gap position and spillover regions in comparison to FGDs created with simple *wh*-phrases (Hofmeister and Sag, 2010). This observation has been interpreted as reflecting the claim that *which*-N fillers have high levels of activation in WM due to their semantic and structural complexity, meaning that they are more easily accessed and retrieved at the integration site (in comparison to simple *wh*-fillers, which have a low level of activation in memory). However, the slowdown in RTs in the current study is not consistent with the assumptions of Hofmeister and Sag's (2010) complexity account: if *ʔayya-še* ('*which*-N') phrases are maintained

actively in WM by the time gaps are processed, while *ʔayš* ('what') phrases are not, a facilitation effect is predicted.¹⁰

A further challenge for the complexity-based processing account is that it predicts a correlation between WM capacity and the acceptability of island-violating dependencies, which was not found to be the case: in the present study, the acceptability of *which*-questions with gaps was not predicted by participants' WM capacities.

In summary, the present results revealed that *ʔayya-še* ('which-N') fillers ameliorate island effects; this advantage in processing is attributed to their high level of activation in memory compared to *ʔayš* ('what') fillers when the integration site is processed. After discussing whether their high level of activation level in memory is triggered by their D-linking properties (Frazier and Clifton, 2002) or by their semantic and structural complexity (Hofmeister and Sag, 2010), we concluded that further experimental investigation is required to fully tease apart these two accounts.

6.7.3 Alternative analysis

In this section, we will consider an alternative interpretation for the interaction between resumption and D-linking as observed in the current study. The contrast between D-linked wh-questions and non-D-linked wh-questions might be attributed to the topic-like nature of D-linked fillers: D-linked wh-phrases (but not no-D-linked ones) contain a lexical element beside the wh-word itself, and their interpretation depends on elements previously mentioned in the discourse.

It has been argued that topics in Arabic, as in Clitic Left Dislocation (CLLD), obligatorily bind pronominal clitics at the internal position where they are semantically interpreted as in (29). Importantly, these structures are immune to island effects as explained in (30).

29. naadia šeeḥ-a saami mbeeriḥ
 Nadia saw.3ms-her Sami yesterday
 'Nadia, Sami saw her yesterday.'

¹⁰ Which-N fillers also elicited longer RTs compared to simple questions in Dutch (Donkers et al., 2013), and German (Freitag et al., 2013; Freitag and Repp, 2015).

(Aoun et al., 2009, p.191)

30. sməʃt ʔənnə naadia rəʔht mən duun ma
 heard.1s that Nadia left.2ms without Comp
 təhke maʃ-a
 talk.2ms with-her

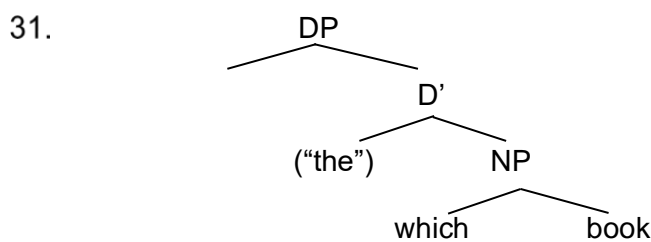
'I heard that Nadia, you left without talking to her.'

(Aoun et al., 2009, p.201)

To the extent that D-linked wh-phrases are interpreted as topics as in CLLD, they are expected (i) to obligatorily bind pronoun clitics inside the clause and (ii) to be insensitive to island effects.

The insensitivity of topics (as in CLLD and D-linked wh-questions) to islands is implemented in syntax in different ways. Beside the base generation account (Aoun et al., 2009), it has been argued that movement is involved in the derivation of these structures (Boeckx et al., 2004). Generally speaking, movement-based accounts assume that the contrast between D-linked and non-D-linked questions, in terms of their (in)sensitivity to islands, is due to the fact that the type of movement that is involved in the derivation of D-linked wh-questions is of a different nature from standard operator-driven movements that is involved in the derivation of non-D-linked wh-questions.

For instance, Boeckx et al. (2004), proposed the following structure for D-linked fillers:



Boeckx et al. (2004) proposed that D-linked wh-questions with RPs are derived by stranding the D-head, which might be realised as an overt resumptive pronoun, and extracting the wh-expressions for 'non-agreement reasons'. Following Rizzi (2001), Boeckx et al. (2004) argued that the wh-expression in D-linked wh-questions targets positions in the left periphery to express discourse properties like topicality (i.e.

TopicP). Boeckx et al. (2004) argued that this type of movement, unlike operator-driven movements in wh-movement, is insensitive to island effects.

In terms of these accounts, the unacceptability of gapped D-linked wh-questions is not due either to violation of island effects (as in bare wh-questions) or violation of subcategorisation requirement (as in gapped *illi*-structures). Instead, the decline in the acceptability of gapped D-linked wh-questions is due to the lack of resumptive pronouns which are responsible for the presuppositionality interpretation of D-linked questions; in other words, the lack of resumptive pronouns might be not consistent with the information structure of the filler phrase.

Other accounts attributed the insensitivity of CLLD to islands to discourse-related factors, but contrary to Boeckx et al. (2004), these accounts did not treat discourse functions either as incorporated in Phrase Structure or as syntactic features (Alexopoulou 1999; Alzayid, 2020). To the extent that D-linked wh-fillers in questions are interpreted as CLLDed elements in CLLD, these accounts would assume that the difference between bare wh-questions and D-linked wh-questions in terms of their (in)sensitivity to islands, as observed in the current study, is due to discursive factors rather than syntactic factors; i.e. irrespective of whether movement is involved in the derivation of these structures or not, the main factor responsible for presence or absence of island effects is discursive in nature; D-linked wh-questions with resumptive pronouns circumvent island effects because of their specific discourse function that is not present in non-D-linked wh-questions.

Under such an approach, the observation that the acceptability of island-violating D-linked wh-questions with gaps are degraded compared to RP-dependencies means that the presence of RPs is necessary so that these questions sound felicitous in discourse; in other words, the lack of resumptive pronouns render these questions infelicitous in discourse.

However, it remains unclear under these accounts why resumptive pronouns are obligatory inside islands and optional in non-island D-linked wh-questions; in other words, it is not clear why D-linked fillers should always be topicalised in island-violating dependencies but not in non-island dependencies.

6.7.4 Task Type

In the current study we found an important impact of the type of task accompanying the self-paced reading experiment on participant behaviour: specifically, the

amelioration effect of resumption and type of filler on island-violating *wh*-dependencies was detected in the self-paced reading with an acceptability judgment task, but not in the self-paced reading with a comprehension question task. The difference between participants' behaviour across the two tasks suggests that participants do not reach the same level of processing in the two tasks. More specifically, it would appear that participants perform the comprehension question version of the experiment with shallow parsing, while deep parsing is involved in the performance of the acceptability judgment version (Stewart et al. (2007)).

We propose that this observation might also explain why the amelioration impact of resumption was not detected in complex *wh*-questions in Study 3, in which the self-paced reading experiment was accompanied by a comprehension question task.

6.8 Conclusion

This chapter has investigated the extent to which RPs inside island-violating D-linked *wh*-questions ameliorate island effects. There are two main views regarding the nature of the interaction between RPs and D-linked fillers: a syntactic view, which is commonly adopted for Arabic, and a processing view, commonly adopted for English. In essence, the syntactic theory assumes that RPs are bound pronominal variables forming part of a syntactic binding dependency in D-linked *wh*-questions. The processing theory, on the other hand, assumes that RPs are interpreted as discourse pronouns to resolve the ungrammatical D-linked FGDs anaphorically.

To assess the predictions made by these theories, we investigated the processing and acceptability of gaps and resumptive pronouns in Baha Arabic using inherently and contextually island-violating D-linked *wh*-questions that were presented in a felicitous context. Context sentences were further used to manipulate whether a referent is salient or non-salient in discourse.

While the two accounts predict that RPs would be preferred to gaps in both types of *wh*-questions, they make different predictions regarding the magnitude of improvement in acceptability caused by RPs (processing: partial improvement; syntactic: high improvement), as well as the sensitivity of RPs to the salience of the referent in discourse (processing: sensitive; syntactic: insensitive).

We found that RPs ameliorate island effects in *ʔayya-še* ('which-N') questions, but not in *ʔayš* ('what') questions. This observation was interpreted as an indication that *what*-questions cannot be contextually D-linked in Baha Arabic.

Our findings further suggest that a distinction must be maintained between RPs in wh-questions and RPs in relative clauses. This is best captured syntactically by assuming that RPs are ‘last resort’ expressions used to fix derivation problems in *which*-questions (Shlonsky, 1992; Aoun, 2000), while constituting part of the initial derivation in relative clauses. This account accordingly assumes that resumption can license grammatical binding dependencies in island-violating wh-questions when the filler phrase is D-linked. A syntactic account of RPs in wh-questions is also compatible with the lack of impact of the discourse salience of their referent.

What remains unaccounted for by such an approach is the variability in the acceptability pattern of island-violating *which*-questions with gaps compared to the consistently rejected island-violating *what*-questions with gaps. This variability indicates that the effect of filler type on islands is not at the grammatical level. We discussed this finding with reference to the assumptions made by the D-linking and complexity-based processing accounts. The insensitivity of parsers to contextually D-linked *what*-questions (if they do D-link), as well as to the salience of referents in discourse, are not consistent with the D-linking account but are rather expected by the complexity-based processing account. However, longer RTs at the integration site in *which*-questions with gaps compared to *what*-questions, as well as the lack of correlation between parsers’ WM capacities and their sensitivity to island effects, are not compatible with the assumptions of the complexity-based processing account.

Moreover, we also found an effect of task type: the amelioration effect of resumption and filler type was observed in the self-paced reading with an acceptability judgment task, but not in the self-paced reading with comprehension question task. We propose that this difference could arise due to the nature of comprehension questions, which encourage shallow rather than deep processing. By contrast, parsers engage in deep processing when the self-paced reading task is accompanied by an acceptability judgment task.

Finally, we concluded that the entire theoretical picture we developed in the present study requires a combination of syntactic and processing factors to be understood; neither a purely syntactic account nor a purely processing account can explain the entire pattern of results. More specifically, we proposed that the amelioration effect of resumption on islands in D-linked wh-questions is at the grammatical level, while that of filler type reflects the impact of extra-grammatical factors.

The implications of our findings across the four studies, as well as suggestions for future work, will be highlighted in the next chapter.

Chapter 7 General Discussion and Conclusion

This dissertation contributes to the debate regarding the nature of factors that underline the distributional properties of resumptive pronouns (RPs) across dependency structures and across languages. The debate particularly centers on whether resumption is a syntactic phenomenon or a processing one and whether this characterization differs across languages (and dependencies). Three alternative possibilities have been proposed in the literature on resumption. One is that resumption is a homogenous, narrow syntax phenomenon in certain languages (possibly restricted to particular dependencies) (i.e., grammatical RPs) while it falls at the interface of syntax and processing (i.e., intrusive RPs) in others (Alexopoulou, 2010; McCloskey, 2017). Alternatively, resumption might be a uniform (syntactic or processing) phenomenon cross-linguistically, operating as a last resort device (i.e., intrusive RPs) (Shlonsky, 1992; Ariel, 1999; Hawkins, 2004). Finally, it is possible that RPs have different functions within the same language, depending on the structure where they occur (non-island vs. island) (Erteschik-Shir, 1992; Sterian, 2016).

In this dissertation, we investigated the phenomenon of resumption in Baha Arabic, a language with a productive use of resumption across different types of dependency structures.

We further investigated the extent to which type of *wh*-filler affect the acceptability and processing of islands in Baha Arabic. This hypothesis has been tested in several languages, including English (Hofmeister and Sag, 2010; Alexopoulou and Keller, 2013; Goodall, 2014), Greek (Alexopoulou and Keller, 2013) and German (Freitag et al., 2013; Freitag and Repp, 2015), but has not been tested in a variety of Arabic. Arabic syntacticians make no such argument for Arabic varieties: the commonplace view is that *which*-questions with violations of island constraints are not acceptable unless a resumptive pronoun appears at dependency tail. To our knowledge, this study is the first one to investigate this hypothesis in a variety of Arabic.

Four experimental studies revealed a rich and complex set of data. In this final chapter, the findings of the four experiments will be reviewed. These findings will in turn enrich the discussion that follows regarding the accounts of processing filler-gap dependencies (FGDs) and the nature of island effects. The cross-linguistic variation in the distributional properties of RPs among varieties of Arabic will be discussed in light of our findings. Suggestions for further research will be finally introduced.

7.1 Dissertation summary:

7.1.1 Summary of the design of the four studies

The following table presents a comparative summary of the design of the four studies.

Table 7-1 Comparative summary of the design of the four experimental studies.

	EXP 1	EXP 2	EXP 3	EXP 4
Task	Offline: acceptability judgment	Online: self-paced reading	Online (self-paced reading)	Online: self-paced reading
	No follow-up interpretation Q	T/F question (general interpretation)	T/F question (general interpretation)	-Comprehension question: answer the wh-question used in the test item -Judge acceptability
	No context	No context	With context (for plausibility)	With referential context
Variables	Tail: Gap vs RP	Tail: Gap vs RP	Tail: Gap vs RP	Tail: Gap vs RP
	Structure: Bare WH/RC/Cleft-WH	Structure: Bare WH/RC	Structure: Bare WH/Which-Qs	Structure: Bare WH/Which-Qs
	Condition: Non-crossed Island Crossed Non-Island	Condition: Non-crossed Island Crossed Non-Island	Condition: Non-crossed Island Crossed Non-Island	Condition: Crossed Adjunct Island

	Crossed Adjunct Island	Crossed Adjunct Island	Crossed Adjunct Island	
	Crossed RC-Island	Crossed RC-Island	Crossed RC-Island	
				Salience: Salient referent (for the RP/gap) / Not salient

Study 1 was designed to define the grammatical status of RPs used in BA *wh*-dependencies. Study 2 aimed to assess the extent to which RPs in *what*-questions have a processing effect and the extent to which online real-time data correlate with offline acceptability data obtained from Study 1. Study 3 further tested the processing advantage of RPs in *wh*-questions, but differs from Study 2 in that (i) *wh*-questions were contextualized to be pragmatically plausible, and (ii) D-linked *which*-questions were tested beside non-D-linked *what*-questions. The observation that resumption is penalized in non-D-linked *what*-questions, but not in D-linked *which*-questions suggested that resumption is sensitive to discourse properties of filler phrases. Thus, Study 4 was designed to test how resumption interacts with discourse linking properties of filler phrase as well as salience of their referents in discourse.

7.1.2 Summary of the main findings

The main findings of the four studies are listed below:

- Interaction between resumption and type of dependency structure:
 - ***III*-structures (i.e. relative clauses and cleft *wh*-questions)**
 - Resumptive pronouns are highly accepted regardless of islandhood.
 - Resumptive pronouns significantly facilitate processing of relative clauses.
 - **Wh-questions:**
 - Interaction between resumption and type of wh-filler phrase:
 - *What-questions:*
 - Non-island:
 - Gaps are significantly preferred, but resumptive pronouns are still accepted marginally.
 - Resumptive pronouns hinder processing, but the processing cost decreases as level of embedding increases.
 - Islands
 - Resumptive pronouns are rated as bad as gaps.
 - Resumptive pronouns neither facilitate nor hinder processing.
 - *Which-questions:*
 - Non-island:
 - Resumptive pronouns neither facilitate nor hinder processing.
 - No acceptability data.

- Island:
 - Resumptive pronouns are highly accepted.
 - Resumptive pronouns significantly facilitate processing.
- Interaction between islands and type of *wh*-filler:
 - Island-violating dependencies with *what*-fillers are consistently rejected.
 - Island-violating dependencies with *which*-fillers received an inconsistent pattern of acceptability with pervasive inter-individual variation.

7.1.3 Interpretation of results

7.1.3.1 Types of RPs in BA

The current dissertation provided evidence that different types of dependencies in Baha Arabic feature different types of RPs. Each dependency structure will be discussed separately below.

7.1.3.1.1 RPs in *illi*-structures

Findings of Study 1 and Study 2 revealed that RPs are obligatory in *illi*-structures regardless of islandhood and that RP-dependencies are processed faster than gap-dependencies. Thus, we argued that resumption in *illi*-structures (relative clauses and cleft *wh*-questions) is a narrow syntax phenomenon, whose distribution is determined by syntactic properties of the relative complementizer (McCloskey, 2006; Alexopoulou, 2010). Resumptive pronouns are lexical entities inserted as part of the numeration before derivation when the morpho-syntactic features of C do not trigger movement, and are interpreted as variables bound by the operator at LF. This is the derivation of RPs in *illi*-structures regardless of islandhood.

7.1.3.1.2 RPs in *wh*-questions

Findings of the four studies suggest that resumption in *wh*-questions is sensitive to type of filler phrase.

Starting with *what*-questions, findings of Studies 1, 2 and 3 revealed that the acceptability and processing of RPs is affected by whether an RP appears inside islands or not. In non-island structures, gap-dependencies in *what*-questions are consistently accepted, but variation is observed in RP-dependencies in Study 1. The online data in Study 2 and Study 3 revealed that non-island RP-dependencies in

what-questions are associated with longer RTs compared to their gapped counterparts at the spillover region (i.e. the region following the integration site), irrespective of whether they were preceded by a context enhancing the plausibility of the sentence.

One possible analysis is that *what*-questions in BA cannot host grammatical RPs: RPs in *what*-questions are not variables but intrusive pronouns with a processing function. The observation that the processing cost of RP-dependencies in *what*-questions decreases as complexity of the dependency increases in Study 2 and Study 3 might be consistent with the intrusive (processing) analysis of these RPs. However, the observation that these RP-dependencies are never processed faster than gap-dependencies, contrary to what happens with intrusive RPs in English (Hofmeister and Norcliffe, 2013; Hammerly, 2019), is not consistent with an intrusive RP analysis. We suggested instead that parsers become less sensitive to the ungrammaticality of RPs in *what*-questions as demands on WM resources increase.

We suggested that the marginal level of acceptability of RPs in *what*-structures observed in Study 1 might be due to the existence of true RPs in *illi*-structures (as a kind of cross-structural influence). However, the lack of processing effect of RPs in complex non-island *what*-questions needs further exploration. Specifically, in the current study, we only tested dependencies with a single level of embedding. It could be that the processing advantage of intrusive RPs is evident in dependencies with more levels of embedding.

An alternative analysis is that, in line with accounts that draw a distinction between RPs in island and non-island contexts, RPs in non-island *what*-questions have an interpretive effect and their interpretation occurs at the syntax-pragmatic interface (Sterian, 2016); i.e. resumptive pronouns are only used to convey a particular interpretive effect (which gap-dependencies cannot convey) but not for a processing function. The lack of contexts that encourage this interpretation might lead participants to be uncertain about their interpretation. According to Sterian (2016), such an interpretive effect is not present when the pronoun occurs inside island structures; instead, RPs inside island structures have an intrusive (syntactic or processing) last resort function.

In non-island *which*-questions, by contrast, there was no significant difference between the processing of RP-dependencies and gap-dependencies in Study 3, suggesting that RPs in *which*-questions have a different status than in *what*-questions, because of the inherent discourse-linking properties of the *which*-filler.

Acceptability judgment data for RPs in non-island *which*-questions will be required to determine whether they function as true RPs or not.¹¹

Moving to island-violating *wh*-questions, we found that these structures are more acceptable and easier to process when they contain a discourse-licensed RP; i.e. in *which*-questions. In island-violating *what*-questions, on the other hand, RPs are consistently rejected, and have no facilitation effect on processing, even in the presence of a discourse licensing context in Study 4. The observation that discourse licensing of RPs is restricted to dependencies with inherently D-linked fillers (*ʔayya-še* 'which-N') suggests that *what*-questions cannot be contextually discourse-linked in BA (at least in island configurations).

We argued that the insensitivity of RPs in island-violating *which*-questions to the manipulation of salience of referents in discourse as well as their relatively high acceptance rate suggest that the amelioration effect of RPs is at the grammatical level, which means that RPs are not interpreted as discourse pronouns but as bound variables similarly to traces.

At the same time, our findings suggest that a distinction must be maintained between RPs in island-violating *which*-questions and RPs in island-violating *illi*-structures, despite their similarities (i.e. both interpreted as syntactically bound variables). In particular, two observations lead us to the conclusion that RPs in *ʔayya-še* 'which-N' questions are not true RPs in binding dependencies as in relative clauses: (i) while gap-dependencies in relative clauses were clearly rejected in Study 1, they received an inconsistent pattern of acceptability, with a pervasive inter-individual variation, in *ʔayya-še* 'which-N' questions in Study 4, and (ii) the processing advantage of RPs in *ʔayya-še* 'which-N' questions are not detected when comprehension questions allow shallow processing in Study 3 and the sub-experiment with Comprehension Question task in Study 4, while it is detected for RPs in relative clauses under similar conditions in Study 2.

Therefore, we propose that RPs in island-violating *wh*-questions are intrusive RPs, whose presence is restricted to *which*-questions, because of their inherent discourse linking properties.

¹¹ Sterian (2016) argued that RPs in non-island D-linked *wh*-questions are discourse pronouns that are utilized for an interpretive, rather than syntactic or processing, function.

We discuss the intrusive nature of these RPs under a syntactic last resort account. The syntactic last resort analysis assumes that RPs can be inserted as last resort expressions to repair violations of syntactic movement (Shlonsky, 1992; Aoun, 2000) when the filler phrase is D-linked. In other words, unlike true RPs in *illi*-structures that constitute part of the initial numeration in the derivation, RPs in *which*-questions are last resort expressions that are used to fix derivation problems when movement is illicit.

Evaluating syntactic last resort account for intrusive RPs

It is commonly argued that the technical implementation of the syntactic last resort account faces several problems and that further illustration is required to account for the syntactic mechanisms involved in the transfer from movement to binding relations during the derivation (Salzmann, 2009; Reitbauer, 2013).

The optionality of RPs in non-island contexts is further viewed as problematic for syntactic last resort accounts of resumption. Indeed, the theoretical literature on resumption in Arabic claims that RPs are optional in non-island contexts in *which*-questions (Aoun et al., 2009; Sterian, 2016), see (1).

1. a. With Gap:
Iman ya: riɖɖʒa:l ša:fet ____ bi-l-ḥafla
Iman which man saw.3FS ____ at-the-party
'Which man did Iman see ____ at the party ?'
b. With RP:
Iman ya: riɖɖʒa:l ša:fet=**eh** bi-l-ḥafla
Iman which man saw.3FS=**3MS** at-the-party
'Which man did Iman see [**him**] at the party?'
(Iraqi Arabic; Sterian, 2016)

Although we don't investigate the acceptability of RPs in non-island *which*-questions in the current study to determine their grammatical status, the online data in Study 3 suggest that RPs are not penalized in non-island contexts in *which*-questions as in non-island contexts in *what*-questions.

However, this pattern of optionality would not be problematic for syntactic last resort accounts for RPs inside islands if a distinction is made between the function of RPs

in non-island and island contexts. This is the view that is adopted for RPs in Iraqi Arabic (Sterian, 2016). For instance, Sterian (2016) argued that optional RPs in non-island D-linked wh-questions have an interpretive effect, as explained in (2).

2. A. Possible answers to questions with quantifiers - gap

ya: mraya kull riḏḏa:l ʕazam____
which woman every man invited.3MS

'Which woman did every man invite ____ ?'

a. Natural function answer: his sister

b. Pair-list answer: Samer, Suha; Ahmad, Najwa; etc

B. Possible answers to questions with quantifiers - resumption

ya: mraya kull riḏḏa:l ʕazam=ha
which woman every man invited.3MS=3FS

'Which woman did every man invite [her] ?'

a. Natural function answer: his sister

*b. Pair-list answer: Samer, Suha; Ahmad, Najwa; etc

(Sterian, 2016, pp.212–213)

Alternatively, it is possible that the optionality between RP-dependencies and gap-dependencies in non-island contexts reflects an alternation between movement and binding dependencies, where this alternation is determined by processing factors. Specifically, as explained in 3.5.1.2.2, we proposed that, following (Alexopoulou, 2010), binding dependencies are less demanding in terms of processing in comparison to movement dependencies, since no fillers are required to be integrated at intermediate CPs (which is the case in movement dependencies).

We conclude that, among the accounts we consider for RPs in BA island-violating *which*-questions, a syntactic last resort account is most consistent with the findings. However, a proper explanation for the mechanisms in which the transfer from movement to binding dependencies takes place in syntax and for whether RPs in non-island contexts have a pragmatic or a processing function in BA is still required to supplement this hypothesis.

7.1.3.2 Interaction between type of *wh*-filler and islands

A further contribution for this dissertation is the investigation of the extent to which type of filler affect the acceptability and processing of islands in BA. To our knowledge, this study is the first one to investigate this hypothesis in a variety of Arabic. Arabic syntacticians commonly argue that island-violating *which*-questions with gaps are not acceptable unless an RP appears at dependency tail. In other words, there is no argument that *which*-fillers ameliorate island effects in Arabic in absence of RPs, as it is the case in other languages, such as English. This dissertation investigated this hypothesis in Study 4 and detected a significant impact of type of filler on acceptability and processing of islands: while island-violating *which*-questions with gaps received an inconsistent pattern of acceptability, island-violating *what*-questions with gaps are consistently rejected. The acceptability pattern of island-violating *which*-questions with gaps is not accountable under a syntactic account, as it is gradient and manifests a great deal of inter-individual variation. Thus, we argued that the impact of *which*-fillers on islands must be treated as an extra-grammatical phenomenon.

We argue that the processing advantage of *which*-fillers is attributed to their high level of activation in memory compared to *ʔayš* 'what' fillers when the integration site is processed. We discussed whether their high level of activation level in memory is triggered by their D-linking properties (Frazier and Clifton, 2002) or by their semantic and structural complexity (Hofmeister and Sag, 2010), and we concluded that, teasing apart between these two accounts requires further experimental investigation.

7.1.4 The effect of the behavioural tasks

The phenomena of resumption and islands in varieties of Arabic have been generally addressed in formal 'theoretical' syntax literature. In the current study, we adopted an experimental approach through the use of formal acceptability judgment tasks and online self-paced reading tasks that enable us to obtain fine-grained observations. For instance, the contrast between the acceptability pattern and the real-time processing of island-violating *which*-questions and *illi*-structures with gaps and RPs gave us insight into the nature of RPs used in each structure, leading us to conclude that these two structures feature two different types of RPs. Additionally, the use of formal acceptability judgment tasks allows us to detect the amelioration effect of type of *wh*-filler on island-violating dependencies in BA. No such claims exist in the

theoretical syntax literature on varieties of Arabic, which is primarily based on informal methods for collecting acceptability judgment data.

Furthermore, we found an important impact of the type of task accompanying the self-paced reading experiment on participants' behavior: specifically, the amelioration effect of resumption and type of filler on island-violating *wh*-dependencies was detected in the self-paced reading with an acceptability judgment task, but not in the self-paced reading with a comprehension question task. The difference between participants' behavior across the two tasks suggests that participants do not reach the same level of processing in the two tasks. More specifically, it would appear that participants perform the comprehension question version of the experiment with shallow parsing, while deep parsing is involved in the performance of the acceptability judgment version (Stewart et al. (2007)).

Our findings also indicate that behavior of participants is affected by whether the experimental sentences are contextualized or not. Although we don't have data that are designed to directly address the relevance of contextualization factor to participants' behavior, we still have an indication that detailed discourse context has an effect. Specifically, comparing the acceptability of island-violating *what*-questions in Study 4 (where they are preceded with detailed contexts) with their acceptability in Study 1 (where they are decontextualized) shows that the probability of rejecting these structures with the lowest acceptability scores decreased from 75% in experiment 1 to 50% in Study 4. This decrease in their rejection rate might reflect the impact of introducing detailed context before experimental sentences, increasing their plausibility.

In addition, when designing the experimental sentences of Study 4, we intended to make them quite plausible through maximizing the degree of cohesion between the matrix clauses and the embedded adjunct clauses. A subset of the experimental sentences in Study 4 presented below; these examples involve a causation relation between the matrix clause and the adjunct clause.

3. ʔayy ʕarḍ / ʔayʃ muna zeʕlet lamman
 Which presentation/what Muba felt-ubse-3SF when
 al-munaẓmeen alḡauu-/h men parnamaj
 the organisers removed_/it from program
 al-muʔtamar?

the conference?

Which presentation/what was Mona upset when the organisers removed _/it from the conference program?

4. ʔayy baħθ/ ʔayš traqqa al-ʔustaq lamman

Which research/what promoted-M.PASS the doctor when

našar-/uh fi al-majallah al-ʕelmyyah?

published _/it in the-journal the-international?

Which research/what was the doctor promoted after he published _/it in an international journal?

5. ʔayy kream/ ʔayš ʔaθar al-ħuruuq ekhtafat

Which ointment/what scars the burn disappear

lamman sara estakhdamat-/uh ʕla yedda-ha al-yesar?

when Sara used -/it on hand-her the-left?

Which ointment/what did the burn scars disappear when Sara used -/it on her left hand?

We speculate that this aspect of experimental sentences enables us to detect a significant and reliable effect of both type of filler and resumption on island-violating dependencies in Study 4.

7.1.5 Conclusion

Unlike previous experimental studies on resumption in grammaticalized resumption languages, which detect a small increase in acceptability of island violating dependencies when an RP appears at dependency tail (Farby et al., 2010; Keshev and Meltzer-Asscher, 2017; Tucker et al., 2019), leading to questioning the nature of the difference between RPs in grammaticalized resumption languages and intrusive resumption languages, the current study detected a significant increase in the acceptability of island violating dependencies when an RP appears at dependency tail. The high acceptance rate of RPs in island-violating dependencies both in *illi*-structures and *which*-questions suggest that the impact is at the grammatical level, and that RPs in these structures are interpreted as bound variables as traces.

The observation that island-violating *which*-questions and island-violating *illi*-structures with gaps and RPs behave differently in several aspects (both in terms of processing and acceptability), lead us to conclude that resumption in Baha Arabic is not a uniform phenomenon, despite the argument that it mainly has a syntactic function; (i) true RPs in *illi*-structures constitute part of the initial derivation in binding dependencies (as morpho-syntactic features of C do not trigger movement) and (ii) intrusive RPs in *which*-questions are utilized as last resort devices to fix derivation problems when movement is illicit.

The amelioration effect of type of filler phrase, on the other hand, is interpreted as reflecting an extra-grammatical phenomenon. Hence, the picture that emerges from the four studies is that the complexity of *wh*-dependencies, due to either the application of grammatical rules or limitations in WM resources, can be ameliorated or avoided altogether either through syntactic mechanisms (i.e. by establishing binding dependencies with true RPs as in *illi*-structures or utilizing intrusive RPs as syntactic last resort devices as in *which*-questions) or with processing mechanisms (i.e. by using complex fillers instead of simple fillers).

This interpretation of *wh*-dependencies with violation of islands would not be obtainable without the use of a combination of formal acceptability judgments and online experimental methods. At the same time, our findings suggest that designing self-paced reading experiments requires the presence of follow-up tasks that encourage deep, not only shallow, processing.

7.2 Consequences for processing accounts of FGDs

There are two dominant accounts of the processing of FGDs, depending on representation of the filler phrase in WM during the processing of the FGD: the maintenance account and the retrieval account. We will show below that our results are best explained by a combination of the two.

7.2.1 Summary of existent processing accounts of FGDs

Under the maintenance account, the filler phrase has to be maintained active in working memory in a particular storage throughout the processing of FGDs, so that it can be directly accessed and retrieved at the integration site (Wanner and Maratsos, 1978; Gibson, 1998; Wagers and Phillips, 2014). The resource limitation theory of island effect is based on the maintenance view of processing FGDs (Kluender and Kutas, 1993b; Hofmeister and Sag, 2010). It argues that parsers'

ability to hold the filler active in WM is negatively affected when demands on WM resources increases, for instance, when processing island clause boundaries. On the other hand, D-linking and/or the semantic complexity of fillers maximizes the likelihood of maintenance in memory, and therefore, ameliorates island effects (Kluender and Kutas, 1993b; Hofmeister and Sag, 2010).

Retrieval accounts, on the other hand, assume that filler phrases are not maintained in a special storage; rather, all the words in the dependency are stored. When the gap is encountered, the filler phrase must be reactivated and retrieved (Lewis and Vasishth, 2005; Van Dyke and McElree, 2006). Thus, this account predicts that the processing cost arises due to retrieval errors or similarity-interference effect. WM-based accounts for island effects from the perspective of retrieval-based account have not been explicitly proposed. The only one attempt to account for island effects from the perspective of retrieval accounts was done by Michel (2014), who presented the similarity-interference account of islands; according to this account, island effects arise since 'Island boundaries contain features that interfere with the retrieval of fillers.' (Michel, 2014, p.69). Semantic complexity of fillers increases the distinctness of features of fillers and therefore decreases inferences errors.

This similarity-interference account of islands cannot be extended to adjunct islands, the type of island we focused on in the current study, as adjunct island boundaries do not share features with filler phrases.

Wagers and Phillips (2014) and Kim et al. (2020) argued that both maintenance and retrieval processes are involved in the processing of FGDs. According to Kim et al. (2020), filler phrases are maintained active in WM (i.e. in a particular storage) throughout the processing of the FGD; however, in certain cases, filler phrases are released from maintenance (i.e. stored where other words of the dependency are stored) and need to be reactivated and retrieved at gap position, as in the second gap in the following structure.

6. Which mistake in the program/programs __ will be disastrous for the company and certainly __ is/are harmful for everyone involved?

In (6), the filler phrase 'which mistake in the program/programs' can be integrated at the gap in the first conjunct, and then, released from maintenance, as the first conjunct can be interpreted as an independent sentence. However, when the connective 'and' is processed, the filler phrase must be reactivated. This is so

because the FGDs in the coordination construction obey Coordinate Structure Constraint and the Across-the-Board movement restriction (Ross, 1967).

Kim et al. (2020) argued, on basis of experimental findings, that (i) maintained fillers (i.e. fillers that are maintained active in memory until the gap is encountered) are less susceptible to decay effects, and thus are easily accessed at integration site, while (ii) unmaintained fillers (i.e. fillers that are released from maintenance before encountering the gap) are subject to decay effects, and thus are harder to access at retrieval sites.

Island effects have not been investigated from the viewpoint of the accounts of processing FGDs that incorporate both maintenance and retrieval processes. We propose the following analysis.

Typically, the active maintenance of filler phrases is sensitive to the complexity of the derivation of FGDs. The observation that active filler strategy is not at work inside island structures (Frazier and d'Arcais, 1989) means that filler phrases are released from maintenance inside island structures. That is, the filler phrase is released from maintenance upon encountering an island boundary due to (i) the application of grammatical constraints that do not allow gaps to occur inside islands, and/or (ii) parsers' inability to maintain the filler as WM resources are overloaded (Phillips, 2013).

Under Kim et al.'s (2020) maintenance/retrieval account, the observation that island-violating dependencies are accepted in certain cases suggests that parsers are able to reactivate and retrieve unmaintained fillers inside islands.

7.2.2 Implication of findings on accounts of processing FGDs

Typically, under the maintenance-based view, the slowdown in RTs inside islands at integration site (compared to non-islands) in argument dependencies is interpreted as reflecting the complexity of processing a verb without an argument, as no maintained filler is available to be reactivated (Hammerly, 2019). Hofmeister and Sag's (2010) finding that *which*-fillers are processed faster than *what*-fillers inside islands in English suggests that it is however possible for parsers to maintain 'prominent' fillers active in memory inside islands. Our findings furthermore suggest that parsers are actually able to interpret the filler phrase at the integration site inside islands as suggested by the offline data of island-violating *which*-questions with gaps (as these sentences are accepted).

However, contrary to Hofmeister and Sag (2010), who interpreted the facilitation of processing island violating dependencies with *which*-fillers compared to *what*-fillers as reflecting the fact that *which*-fillers remains active in WM by the time integration site is processed, we found that the processing of *which*-questions is slower than *what*-questions at integration site inside islands in Baha Arabic. This finding suggest that (i) *which*-fillers are not maintained active inside islands (contrary to maintenance-based view) and that (ii) parsers are able to reactivate ‘unmaintained’ fillers when the integration site is encountered (as suggested by the offline data), in line with Kim et al. (2020).¹²

Thus, we propose that, although maintenance fails inside islands, encountering the gap triggers a retrieval process that targets the ‘unmaintained’ filler phrase (gap antecedent). Since unmaintained fillers are subject to decay effect, retrieval process is complex and might be not successful (Kim et al., 2020). However, *which*-fillers have an advantage in terms of resistance to memory decay effect, compared to *what*-fillers, when both are unmaintained inside islands, due to the availability of more lexical and semantic cues over retrieval (Hofmeister and Sag, 2010) and due to their inherent D-linking properties (Kluender and Kutas, 1993b; Kluender, 1998). In *what*-questions, on the other hand, parsers process the verb with no argument, as the retrieval process is likely to fail due to the low level of activation of non-complex fillers. We argue that this analysis explains why island-violating dependencies with *which*-fillers are accepted more than island-violating dependencies with *what*.

However, re-accessing the content of unmaintained *which*-fillers and retrieving their rich semantic and lexical content, be it successful or not, is costly (Freitag and Repp, 2015). This fact, in line with the violation of grammatical constraints on movement, explains the inconsistent pattern of acceptability as well as the slowdown in RTs of *which*-questions with gaps at spillover region compared to *what*-questions: when the retrieval process is not successful, *which*-questions with gaps are not accepted, and when the retrieval process is successful, *which*-questions with gaps are accepted.

We did not observe a significant correlation between participants’ acceptability of island-violating dependencies and their WM capacities. However, as noted by

¹² The contrast between English data (Hofmeister and Sag, 2010) and our data in regard with processing *wh*-questions inside islands is puzzling, and more exploration is needed to address the source of this distinction.

(Pañeda et al., 2020), the correlation between individuals' WM capacity and their acceptability or processing of island-violating dependencies might not be a reliable measure to test the predictions of WM-based accounts. It is possible that the type of task we use does not tap at the part of memory that is involved in the processing of FGDs (Hofmeister et al., 2012). Measures of WM capacity might be predictive of information maintenance, but we have argued that maintenance-only accounts are insufficient to account for the processing of FGDs (as retrieval is another important dimension). The cognitive cost of retrieving the lexical content of the filler phrase might in part be due to retrieval errors or similarity-interference effect. Future research will be required to identify the optimal cognitive predictors in that respect.

We assumed in 7.1 that the availability of the syntactic last resort strategy in *which*-questions, but not in *what*-questions, is due to the inherent D-linking properties of *which*-fillers. In light of the above discussion regarding the mechanisms involved in processing FGDs, we propose an alternative explanation for the contrast between *which*-questions and *what*-questions in terms of their interaction with resumption. Specifically, it could be that RPs can be used as syntactic last resort devices only when filler phrases are prominent in memory when the integration site is processed. On the other hand, RPs cannot be used as last resort devices to establish binding dependencies in case of *what*-questions due to the absence of an accessible filler that can be potentially retrieved from memory by the time RPs are processed. In other words, RPs cannot be interpreted as syntactic variables bound by the filler phrase unless the filler phrase is accessible for retrieval from WM by the time integration site is processed.

This explanation is further consistent with the observation that RPs are preferred with animate fillers (i.e. *who*) compared to inanimate fillers (i.e. *what*) in English and Greek (Alexopoulou and Keller, 2013), Arabic varieties (Aoun et al., 2009) and Irish (McCloskey, 2017). Animate fillers are more salient in memory than inanimate fillers (Alexopoulou and Keller, 2013). Future research will be required to determine whether animacy affects the activation levels of referents.

7.3 Consequences for accounts of island phenomena

It has been much debated whether island effects arise as a result of violating grammatical constraints or as a result of processing limitations. The current study did not aim to address this debate directly. However, the interaction of resumption with islands across the different types of dependency structures tested in the current

study, the impact of type of filler on islands, as well as the direct comparisons between the acceptability data and the online processing data bring additional insights into the nature of island effects.

The most salient finding in the current study is the distinct behavior of *illi*-structures and *wh*-questions in terms of their sensitivity to resumption and islands. This finding is viewed as absolute evidence that the underlying mechanisms for establishing A-bar dependencies in BA are not identical, supporting the syntactic approach to islands. Specifically, the presence of island-effects in *what*-questions where RPs are not allowed, and the absence of such an effect in *illi*-structures where RPs are obligatory (in Study 1), as well as the distinct pattern of the acceptability and processing of island violating *which*-questions and relative clauses with and without RPs (across Studies 2,3,4), support the grammatical accounts' claim that there are two mechanisms for encoding syntactic dependencies (i.e., movement-relations which are sensitive to island-effects (*wh*-questions) and binding-relations that are insensitive to island-effects (*illi*-structures)). The WM-based accounts, on their own, cannot explain these findings.

Nevertheless, results of Study 1 indicate that processing factors contribute to island effects too. Specifically, two observations are not accountable under the syntactic accounts. RP-dependencies crossing islands in *illi*-structures received lower acceptance rate in comparison to their counterparts where no islands are crossed. Due to the modest impact of islands on the acceptability of *illi*-structures, the decline in acceptability is best interpreted in terms of processing rather than the violation of a syntactic constraint. As for the decline in the acceptability of non-island longer dependencies compared to shorter ones in *wh*-questions, it is only accountable under WM-based accounts.

The contribution of the processing factors to islands was revealed most clearly by the acceptability and online processing data in Study 4. The acceptability judgment data revealed a gradient/inconsistent pattern of acceptability of island-violating *which*-questions. At the same time, *what*-questions with gaps were highly rejected. This pattern of acceptability is not predicted by grammatical accounts, as it suggests that, in line with processing accounts, island-violating dependencies are ameliorated when filler phrases are robustly represented in memory (due to their D-linking properties or the complexity of their lexical and semantic content).

At the same time, the high acceptance rate of RPs in island-violating *which*-questions, in line with their insensitivity to discourse salience, is consistent with syntactic accounts of both resumption and islands.

Hence, even if it is concluded that syntactic constraints are substantial in accounting for island effects, this should not be the final conclusion: there are processing factors at play too. Neither the syntactic, nor the processing factors, would account for the entire pattern of the acceptability data on their own.

Beside, we believe that increasing plausibility of experimental sentences helped to decrease island effects in *which*-questions in Study 4. It is well known that plausibility affects the interpretation of sentences in real time processing (Traxler and Pickering, 1996). In Study 4, the plausibility of experimental sentences was maximized through the introduction of a relevance-maximising discourse context before each experimental sentence. We propose that this might positively affect how participants interpreted island-violating dependencies in real time, leading to a higher acceptance rate for these structures, as explained in 7.1.4.

In addition, when designing the experimental sentences of Study 4, we intended to make them quite plausible through maximizing the degree of cohesion between the matrix clauses and the embedded adjunct clauses, as explained in 7.1.4. It has been argued that lack of cohesion between propositions in island-violating dependencies constitutes a significant factor in determining their acceptability (Kehler, 2002; Dawei Jin, 2016; Chaves and Putnam, 2020). Chaves and Putnam (2020) and Dawei Jin (2016) proposed that the interpretation of sentences with violations of adjunct islands is sensitive to the degree of semantic/pragmatic cohesion between matrix clauses and adjunct clauses: extraction is tolerated when the relation between the propositions is coherent and the extracted element remains salient through the scenario of the utterance.

For instance, Dawei Jin (2016) argued that extraction is degraded in (7) since it describes two parallel and separate actions. However, in (8), the matrix and adjunct clause stand in a causation relation: the second event is construed as a consequence of the first event. Thus, the extracted element remains salient throughout the utterance scenario.

7. *What_i did John build this house [thinking about t_i]?

(Dawei Jin, 2016, p.34)

8. What_i did Mary drive John crazy [trying to fix t_j]?

(Dawei Jin, 2016, p.33)

This view is closely related to the argument that islands arise when the extracted element is presupposed (not part of the pragmatic focus) (Erteschik-Shir, 1973; Goldberg, 2013). Specifically, as adjunct islands are modifiers by nature (i.e. not selected by a head), Chaves and Putnam (2020) proposed that the more these modifiers are incidental, the less likely they are part of the pragmatic focus.

It is therefore plausible to assume that participants rated island-violating *which-N* questions with gaps as highly acceptable when the coherence between the matrix clause and adjunct island clause is maximized, and rated them as unacceptable (or in between) when the degree of coherence between the matrix clause and adjunct island is low.

Although this view seems promising, our results suggest that the inconsistent pattern of acceptability in *which-N* questions with gaps is not conditioned by items (see Figure 6-3); i.e. no item has been consistently rated either as highly acceptable or as highly unacceptable. If degree of coherence between the adjunct clause and the matrix clause has a significant impact on acceptability, items with high degree of coherence would be consistently rated highly acceptable, but this was not the case. However, as we do not have an item analysis following an objective measure of cohesion, we cannot assume that this had no impact.

Comparing our results with that of Tucker et al.'s (2019) experimental study on MSA indicate that maximizing plausibility of experimental sentences (through introducing detailed contexts and maximizing semantic/pragmatic cohesion between propositions of the utterance) might have an impact on acceptability of island-violating dependencies. Particularly, *which*-questions with gaps and RPs in Tucker et al.'s (2019) study on MSA received very low acceptability ratings, despite the theoretical literature's claim that resumption is grammatical in island-violating *which*-questions in MSA. We speculate that Tucker et al. (2019) couldn't pick up the effect of D-linking both on resumption and islands because plausibility was too low in the first place.¹³

¹³ Results of Tucker et al.'s (2019) experimental study might be affected by (im)plausibility of experimental sentences: they were decontextualized, and degree of coherence between propositions in

We hypothesise that the highest level of acceptability for island-violating dependencies with gaps is obtained when a **high level of cohesion** combines with **successful maintenance** of fillers in WM, which is afforded by the presence of D-linked/complex fillers. When these factors combine with **resumption**, the highest acceptance rate obtains. Consequently, we believe that island phenomenon is a multifactorial one, supporting Chaves and Putnam's (2020) eclectic view of islands.

Island phenomena have been central topic of many studies that addressed the accessibility of L2 learners to Universal Grammar (Omaki and Schulz, 2011; Saad Aldosari, 2015; Kim and Goodall, 2016; Perpiñán, 2020, among others). However, the present study questions the argument that islandhood is a purely syntactic phenomenon, suggesting that using them to test L2 accessibility to UG in certain context is a questionable approach. Nevertheless, we propose that resumption seems to be an ideal environment to address accessibility of L2 learners to UG, as our findings suggest that they represent a purely syntactic phenomenon in BA (and possibly in other varieties of Arabic), at least in *illi*-structures, in contrast to languages like English.

7.4 Variation across Arabic varieties

It is commonly argued in the theoretical literature that resumption is obligatory in *illi*-structures in a number of varieties of Arabic, for instance, in Palestinian Arabic (Shlonsky, 1992), Lebanese Arabic (Aoun et al., 2009) and Egyptian Arabic (Choueiri, 2017). Our findings are consistent with this claim: RPs are obligatory in *illi*-structures regardless of islands in BA. However, it does appear that gaps *are* allowed in *illi*-structures in some varieties of Arabic: for instance, according to Choueiri (2017), Moroccan Arabic allows gaps and RPs in non-island relative clauses, as in (9).

9. žbar-t l-ktaab lli nsiti-(h) f-l-qism

island-violating dependencies might not be taken into consideration when designing (some of) the experimental sentences, as indicated by the following examples.

- (i) Which long speech do you feel happy if the king wrote yesterday?
- (ii) Which painting do you blush if the artist painted today?

There is an added problem of using MSA to investigate the acceptability of resumption, as there are no native speakers of MSA.

found-1.sg. the-book that forgot.2.sg.-(it) in-the-class
'I found the book that you forgot in the classroom.'

(Moroccan Arabic; Choueiri, 2017, p.135)

Recently, Al-Aqarbeh and Sprouse's (n.d.) found, in an experimental study on resumption in Jordanian Arabic, that some participants accepted gaps in non-island relative clauses (which involve *illi* complementizer), despite the evident preference of RPs. Inside islands, they found that RPs restore island-violating relative clauses to full acceptability only when they occur inside adjunct and whether islands. No amelioration effect was observed when RPs occur inside RC island structures, as observed in the current study. Our findings contrast with those Al-Aqarbeh and Sprouse's (n.d.) in that (i) gap-dependencies are consistently rejected in relative clauses in BA and (ii) RPs restore relative clauses with violation of RC-islands into full acceptability in BA. This pattern of contrast between resumption in *illi*-structures across varieties of Arabic is not expected under the complementizer-type syntactic account.¹⁴ Further research is clearly needed to inform the systematic study of resumption in *illi*-structures across different varieties of Arabic and across different types of islands.

In *wh*-questions, the distribution of RPs seems to be conditioned by the type of filler phrase across different varieties of Arabic. The theoretical literature suggests that the majority of Arabic varieties allow resumption in *which*-questions (Aoun et al., 2009). Our findings are consistent with this claim. The picture is complicated in case of *what*-questions. It is commonly argued that varieties of Arabic disallow resumption in *what*-questions (Aoun et al., 2009). Tucker et al. (2019) found that RPs are highly penalized in *what*-questions in MSA. In the current study, we found that BA speakers marginally accepted resumption in non-islands *ʔayš* 'what' *wh*-questions, but rejected it inside islands. Similar observations are found in Al-Aqarbeh and Sprouse's (n.d.) experimental study on resumption in Jordanian Arabic, where RPs are accepted marginally in non-island *ʔayš* 'what' *wh*-questions, but are rejected inside islands. Interestingly, BA and Jordanian Arabic use the same form of 'what' filler phrase (i.e. *ʔayš*), while MSA use the word *maḡaa* 'what'. Abdel Razaq (2011) claimed that these two forms of 'what' features distinct morpho-syntactic features. Further exploration

¹⁴ The modality of the acceptability judgment task might have an effect here: Al-Aqarbeh and Sprouse's (n.d.) acceptability judgment experiment was an auditory one while ours was a written one.

is required to uncover the nature of the factors underlying the interaction between type of fillers and resumption in *what*-questions in Arabic varieties.

7.5 Future work

Firstly, the fact that *illi*-structures (binding dependencies) share analogous sentence processing mechanisms with *wh*-questions (movement dependencies) but are still insensitive to island effects means that *illi*-structures can be used as a baseline to investigate the processing mechanisms involved in processing island-violating dependencies.

Second, since the impact of animate fillers on resumption has been paralleled with that of D-linked fillers in Arabic (Aoun et al., 2009), English and Greek (Alexopoulou and Keller, 2013) and Irish (McCloskey, 2017), it would be worth investigating the role of animacy of fillers on acceptability and processing of islands and resumption in Arabic, and show how it relates to the processing and syntactic accounts we addressed in the current study.

The status of RPs in non-island *wh*-questions, furthermore, needs to be investigated to ascertain the extent to which they represent a case of optionality, and if that is the case, what the nature of the factors that underline their optionality, and how individual variation interacts with optionality. Furthermore, it will be interesting to investigate how the acceptability of RPs in non-island *what*-questions across varieties of Arabic is affected by the different morpho-syntactic features of 'what' filler phrases.

It would also be worth investigating the source of the observed contrast between the processing of island-violating dependencies with *what/which* fillers in English (Hofmeister and Sag, 2010), on one side, and Baha Arabic and German (Freitag et al., 2013; Freitag and Repp, 2015) on the other side: despite the observation that *which*-fillers always improve acceptability of island-violating dependencies, *which*-fillers are found to facilitate processing in the former, but hinder processing in the latter. We suggest that investigating the differences in processing island-violating *wh*-questions across languages might shed light on understanding the differences in the mechanisms involved in the real time processing of *wh*-questions cross-linguistically.

References

- Abdel Razaq, I. 2011. *Who is what and what is who: the morpho-syntax of Arabic WH*. Ph.D. thesis, Queen Mary, University of London.
- Ackerman, L., Frazier, M. and Yoshida, M. 2018. Resumptive pronouns ameliorate island violations in forced-choice tasks. *Linguistic Inquiry*. **49**(4), pp.847–859.
- Al-Aqarbeh, R. and Sprouse, J. n.d. *Island effects and resumption in spoken Jordanian Arabic: an auditory acceptability judgment study*. [Online]. [Manuscript]. Available from: <https://sprouse.uconn.edu/papers/Aqarbeh%20and%20Sprouse%20-%20resumption.pdf>.
- Aldosari, S. 2015. *The Role of Individual Differences in the Acceptability of Island Violations in Native and Non-native Speakers*. Ph.D. thesis, University of Kansas.
- Alexopoulou, T. 1999. *The syntax of discourse functions in Greek: a non-configurational approach*. Ph.D. thesis, University of Edinburgh.
- Alexopoulou, T. 2010. Truly intrusive: Resumptive pronominals in questions and relative clauses. *Lingua*. **120**(3), pp.485–505.
- Alexopoulou, T. and Keller, F. 2007. Locality, cyclicity, and resumption: at the interface between the grammar and the human sentence processor. *Language*. **83**(1), pp.110–160.
- Alexopoulou, T. and Keller, F. 2013. What vs. who and which: Kind-denoting fillers and the complexity of whether-islands. In: J. Sprouse and N. Hornstein, eds. *Experimental Syntax and Island Effects*. Cambridge: Cambridge University Press, pp.310–340.
- Alshaalan, Y. and Abels, K. 2020. Resumption as a sluicing source in Saudi Arabic: Evidence from sluicing with prepositional phrases. *Glossa: a journal of general linguistics*. **5**(1), pp.1–36.
- Alzayid, A. 2020. *Arabic dislocation*. Ph.D. thesis, University of Edinburgh.
- Aoun, J. 2000. Resumption and last resort. *DELTA: Documentação de Estudos em Lingüística Teórica e Aplicada*. **16**(spe), pp.13–43.

- Aoun, J. and Benmamoun, E. 1998. Minimality, reconstruction and PF movement. *Linguistic inquiry*. **29**, pp.569–597.
- Aoun, J. and Choueiri, L. 1999. Modes of interrogation. *In*: E. Benmamoun, ed. *Perspectives on Arabic Linguistics*. Amsterdam: John Benjamins, pp.7–26.
- Aoun, J., Choueiri, L. and Hornstein, N. 2001. Resumption, movement, and derivational economy. *Linguistic inquiry*. **32**(3), pp.371–403.
- Aoun, J.E., Benmamoun, E. and Choueiri, L. 2009. *The syntax of Arabic*. Cambridge: Cambridge University Press.
- Ariel, M. 1990. *Accessing noun-phrase antecedents*. London: Routledge.
- Ariel, M. 1999. Cognitive universals and linguistic conventions: the case of resumptive pronouns. *Studies in Language*. **23**(2), pp.217–269.
- Asudeh, A. 2004. *Resumption as resource management*. Ph.D. thesis, Stanford University.
- Asudeh, A. 2012. *The logic of pronominal resumption*. Oxford: Oxford University Press.
- Avrutin, S. 2000. Comprehension of discourse-linked and non-discourse-linked questions by children and Broca's aphasics. *In*: L. Y. Grodzinsky, Shapiro and D. Swinney, eds. *Language and the brain: Representation and processing*. San Diego, CA: Academic Press, pp.295–312.
- Beltrama, A. and Xiang, M. 2016. Unacceptable but comprehensible: the facilitation effect of resumptive pronouns. *Glossa: a journal of general linguistics*. **1**(1), p.29.
- Berwick, R.C. and Weinberg, A.S. 1986. *The grammatical basis of linguistic performance: Language use and acquisition*. Cambridge: MIT Press.
- Blythe, H. 2016. Resumption in English: An investigation of usage and acceptability. *Cambridge Occasional papers in Linguistics*. **9**, pp.156–183.
- Boeckx, C. 2003. *Islands and chains: resumption as derivational residue*. Amsterdam: John Benjamins.
- Boeckx, C. and Grohmann, K.K. 2004. Submove: Towards a unified account of scrambling and D-linking. *In*: D. Adger, C. De Cat and G. Tsoulas, eds. *Peripheries*. Dordrecht: Kluwer, pp.241–257.

- Borer, H. 1984. Restrictive relatives in modern Hebrew. *Natural Language and Linguistic Theory*. **2**, pp.219–260.
- Boxell, O. 2012. Discourse-linking and long-distance syntactic dependency formation in real-time. *In: Proceedings of ConSOLE XIX.*, pp.151–175.
- Chacón, D.A. 2015. *Comparative Psychosyntax*. Ph.D. thesis, University of Maryland.
- Chacón, D.A. 2018. How to make a pronoun resumptive. *In: Cascadilla Proceedings Project.*, pp.99–108.
- Chacón, D.A. 2019. Minding the gap?: Mechanisms underlying resumption in English. *Glossa: a journal of general linguistics*. **4**(1), p.68.
- Chacón, D.A., Imtiaz, M., Dasgupta, S., Murshed, S.M., Dan, M. and Phillips, C. 2016. Locality and word order in active dependency formation in Bangla. *Frontiers in psychology*. **7**, p.1235.
- Chao, W. and Sells, P. 1983. On the interpretation of resumptive pronouns. *In: The Proceedings of NELS 13.*, pp.47–61.
- Chaves, R.P. and Putnam, M.T. 2020. *Un-bounded Dependency Constructions: theoretical and experimental perspectives*. Oxford: Oxford University Press.
- Chen, E., Gibson, E. and Wolf, F. 2005. Online syntactic storage costs in sentence comprehension. *Journal of Memory and Language*. **52**(1), pp.144–169.
- Chomsky, N. 1965. *Aspects of the theory of syntax*. MIT Press.
- Chomsky, N. 1986. *Barriers*. Cambridge: MIT Press.
- Chomsky, N. 1973. Conditions on transformations. *In: S. R. Anderson and Paul Kiparsky, eds. A Festschrift for Morris Halle*. New York: Rinehart and Winston, pp.232–286.
- Chomsky, N. 2001. Derivation by phase. *In: M. Kenstowicz and K. Hale, eds. A Life in Language*. MIT Press, pp.1–52.
- Chomsky, N. 1977. On wh-movement. *In: P. Culicover, A. Akmajian and T. Wasow, eds. Formal syntax*. New York: Academic Press, pp.71–132.
- Chomsky, N. 1995. *The minimalist program*. MIT Press.
- Choueiri, L. 2017. Resumption in varieties of Arabic. *In: E. Benmamoun and R. Bassiouney, eds. The Routledge Handbook of Arabic Linguistics*. Abingdon: Routledge, pp.131–154.

- Christensen, R.H.B. 2015. Ordinal Regression models for ordinal data, R package version 2015.1-21. Available from: <http://www.cran.r-project.org/package=ordinal/>.
- Chung, S. 1994. 'Wh-agreement and' referentiality" in Chamorro.". *Linguistic inquiry*. **25**(1), pp.1–44.
- Cinque, G. 1990. *Types of \bar{A} -dependencies*. Cambridge, MA: MIT press.
- Clifton, C. and Ferreira, F. 1987. Discourse structure and anaphora; Some experimental results. In: M. Coltheart, ed. *Attention and performance XXI*. Hillsdale, NJ: Erlbaum.
- Clifton, C. and Frazier, L. 1989. Comprehending sentences with long-distance dependencies. In: G. N. Carlson and M. K. Tanenhaus, eds. *Linguistic structure in language processing. Studies in Theoretical Psycholinguistics*. Dordrecht: Springer, pp.273–317.
- Cowart, W. 1997. *Experimental Syntax: Applying Objective Methods to Sentence Judgments*. Thousand Oaks, CA: Sage Publications.
- Cowles, H.W., Walenski, M. and Kluender, R. 2007. Linguistic and cognitive prominence in anaphor resolution: Topic, contrastive focus and pronouns. *Topoi*. **26**(1), pp.3–18.
- Dawei Jin 2016. *The Semantics-Pragmatics Interface and Island Constraints in Chinese*. Ph.D. thesis, University at Buffalo.
- Deane, P. 1991. Limits to attention: A cognitive theory of island phenomena. *Cognitive Linguistics*. **2**(1), pp.1–64.
- Dickey, M.W. 1996. Constraints on the sentence processor and the distribution of resumptive pronouns. In: M. W. Dickey and S. Tunstall, eds. *Linguistics in the Laboratory*. Amherst, MA: University of Massachusetts, pp.157–192.
- Donkers, J., Hoeks, J.C.J. and Stowe, L.A. 2013. D-linking or set-restriction? Processing Which-questions in Dutch. *Language and Cognitive Processes*. **28**(1–2), pp.9–28.
- Doron, E. 1982. On the syntax and semantics of resumptive pronouns. *Texas linguistics forum*. **19**, pp.1–48.
- Van Dyke, J. and McElree, B. 2006. Retrieval interference in sentence comprehension. *Journal of Memory and Language*. **55**(2), pp.157–166.

- Erteschik-Shir, N. 1973. *On the nature of island constraints*. Ph.D. thesis, Massachusetts Institute of Technology.
- Erteschik-Shir, N. 1992. Resumptive pronouns in islands. *In: H. Goodluck and M. Rochemont, eds. Island Constraints. Studies in Theoretical Psycholinguistics*. Springer, Dordrecht, pp.89–108.
- Erteschik-Shir, N. and Lappin, S. 1979. Dominance and the functional explanation of island phenomena. *Theoretical linguistics*. **6**(1–3), pp.41–86.
- Farby, S., Danon, G., Walters, J. and Shachar, M. Ben. 2010. The acceptability of resumptive pronouns in Hebrew. *In: Proceedings of IATL [Online].*, pp.1–16. Available from: http://linguistics.huji.ac.il/IATL/26/Farby_et_al.pdf.
- Farra, C.D. 2019. Variability in adjunct islands : An experimental study in Italian. *In: 36th West Coast Conference on Formal Linguistics*. Cascadilla Proceedings Project, pp.115–120.
- Fasiolo, M., Goude, Y., Nedellec, R. and Wood, S.N. 2017. Fast calibrated additive quantile regression. *Journal of the American Statistical Association.*, pp.1–11.
- Fassi-Fehri, A. 1993. *Issues in the Structure of Arabic Clauses and Words*. Dordrecht: Kluwer.
- Fedele, E. and Kaiser, E. 2014. Looking back and looking forward: Anaphora and cataphora in Italian. *University of Pennsylvania Working Papers in Linguistics*. **20**, pp.81–90.
- Francom, J. 2009. *Experimental Syntax: Exploring the effect of repeated exposure to anomalous syntactic structure- evidence from rating and reading tasks*. Ph.D. thesis, University of Arizona.
- Frazier, L. and Clifton, C. 2000. On bound variable interpretations: The LF-only hypothesis. *Journal of Psycholinguistic Research*. **29**(2), pp.125–140.
- Frazier, L. and Clifton, C. 2002. Processing ' d-linked ' phrases. *Journal of Psycholinguistic Research*. **31**(6), pp.633–660.
- Frazier, L. and Clifton, C. 1989. Successive cyclicity in the grammar and the parser. *Language and cognitive processes*. **4**(2), pp.93–126.
- Frazier, L. and d'Arcais, G.B.F. 1989. Filler driven parsing: A study of gap filling in Dutch. *Journal of memory and language*. **28**(3), pp.331–344.
- Freitag, C. and Repp, S. 2015. *Filler complexity in filler-gap dependencies: wh-*

- extraction vs. topicalization*. [Online]. [Poster]. Exhibited at the 28th Annual CUNY Conference on Human Sentence Processing. Available from: http://constantinfreitag.de/presentations/Freitag_Repp_2015_Filler_complexity_in_filler-gap_dependencies-wh-extraction_vs_topicalization.pdf.
- Freitag, C., Repp, S., Berlin, H. and Introduction, I. 2013. *Filler complexity in long-distance wh-extractions in German* [Online]. [Poster]. Exhibited at the 19th AMLaP Conference. Available from: http://constantinfreitag.de/presentations/Freitag_Repp_2013_Filler_Complexity_AMLaP_A4.pdf.
- Friedmann, N. and Costa, J. 2011. Last resort and no resort: Resumptive pronouns in Hebrew and Palestinian. In: A. Rouveret, ed. *Resumptive pronouns at the interfaces.*, pp.223–239.
- Gad, R. 2011. *A syntactic study of wh-movement in Egyptian Arabic within the minimalist program*. Ph.D. thesis, University of Leeds.
- Germine, L., Nakayama, K., Duchaine, B., Chabris, C., Chatterjee, G. and Wilmer, J. 2012. Is the web as good as the lab? Comparable performance from web and lab in cognitive/perceptual experiments. *Psychonomic Bulletin & Review*. **19**(5), pp.847–857.
- Gibson, E. 1998. Linguistic complexity: Locality of syntactic dependencies. *Cognition*. **68**(68), pp.1–76.
- Goldberg, A. 2006. *Constructions at work: The nature of generalization in language*. Oxford ; New York: Oxford University Press.
- Goldberg, A.E. 2013. Backgrounded constituents cannot be “extracted” In: F. J. Sprouse, N. Hornstein and B. Dillon, eds. *Experimental syntax and Island effects*. Cambridge University Press, pp.221–239.
- Goodall, G. 2017. Referentiality and resumption in wh-dependencies. In: Jason Strove, Ruth Kramer and Joseph Sabbagh, eds. *Asking the right questions: Essays in honor of Sandra Chung* [Online]., pp.65–80. Available from: <https://escholarship.org/content/qt8255v8sc/qt8255v8sc.pdf#page=73>.
- Goodall, G. 2014. The D-linking effect on extraction from islands and non-islands. *Frontiers in Psychology*. **5**(OCT), pp.1–11.
- Guilliot, N. and Malkawi, N. 2009. When movement fails to reconstruct. In: A. G. Brucart and J. Sola, eds. *Merging Features.*, pp.159–174.

- Haegeman, L. 1994. *Introduction to government and binding theory*. Oxford: Blackwell.
- Hammerly, C. 2019. *The pronoun which comprehenders who process it in islands derive a benefit*. [Online]. [Manuscript]. Available from: <https://ling.auf.net/lingbuzz/004846>.
- Hawkins, J.A. 2004. *Efficiency and complexity in grammars*. Cambridge: Oxford University Press.
- Hedding, A. 2014. *Resumption in Somali: A phase-based account*. Unpublished.
- Heestand, D., Xiang, M. and Polinsky, M. 2011. Resumption still does not rescue Islands. *Linguistic Inquiry*. **42**(1), pp.138–152.
- Hofmeister, P. 2011. Representational complexity and memory retrieval in language comprehension. *Language and Cognitive Processes*. **26**(3), pp.376–405.
- Hofmeister, P. 2007. *Representational complexity and memory retrieval in language comprehension*. Ph.D. thesis, Stanford University.
- Hofmeister, P., Casasanto, L.S. and Sag, I.. 2012. How do individual cognitive differences relate to acceptability judgments? A reply to Sprouse, Wagers, and Phillips. *Language*., pp.390–400.
- Hofmeister, P., Jaeger, T.F., Sag, I.A., Arnon, I. and Snider, N. 2007. Locality and accessibility in wh-questions. *Roots: Linguistics in search of its evidential base*., pp.185–206.
- Hofmeister, P. and Norcliffe, E. 2013. Does resumption facilitate sentence comprehension? In: P. Hofmeister and Elizabeth Norcliffe, eds. *The Core and the Periphery: Data-driven Perspectives on Syntax Inspired by Ian A. Sag. CSLI Lecture Notes*. Stanford, CA: CSLI Publications, pp.225–246.
- Hofmeister, P. and Sag, I. 2010. Cognitive constraints and island effects. *Language*. **86**(2), pp.366–415.
- Huang, C.J. 1982. Move WH in a language without WH movement. *The linguistic review*., **1**(4), pp.369-416.
- Just, M.A. and Carpenter, P.A. 1992. A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*. **99**, pp.122– 149.
- Kallulli, D. 2009. Why and how to distinguish between pro and trace In: *STiL: Proceedings of Incontro di Grammatica Generativa*., pp.206–218.

- Kaye, A.S. 1970. Modern Standard Arabic and the colloquials. *Lingua*. **24**, pp.374–391.
- Kazanina, N., Lau, E., Lieberman, M., Yoshida, M. and Phillips, C. 2007. The effect of syntactic constraints on the processing of backwards anaphora. *Journal of Memory and Language*. **56**(3), pp.384–409.
- Keffala, B. 2011. Resumption and gaps in English relative clauses: Relative acceptability creates an illusion of ‘saving’. In: *Annual Meeting of the Berkeley Linguistics Society* [Online]., pp.140–154. Available from: <http://journals.linguisticsociety.org/proceedings/index.php/BLS/article/viewFile/846/629>.
- Kehler, A. 2002. *Coherence, reference, and the Theory of Grammar*. Stanford, CA: CSLI Publications.
- Keller, F. 2000. *Gradience in grammar: Experimental and computational aspects of degrees of grammaticality*. Ph.D. thesis, University of Edinburgh.
- Keshev, M. and Meltzer-Asscher, A. 2019. A processing-based account of subliminal wh-island effects. *Natural Language and Linguistic Theory*. **37**(2), pp.621–657.
- Keshev, M. and Meltzer-Asscher, A. 2017. Active dependency formation in islands: How grammatical resumption affects sentence processing. *Language*. **93**(3), pp.549–568.
- Kim, B. and Goodall, G. 2016. Islands and Non-islands in Native and Heritage Korean. *Frontiers in Psychology*. **7**(February), p.134.
- Kim, N., Brehm, L., Sturt, P. and Yoshida, M. 2020. How long can you hold the filler: maintenance and retrieval. *Language, Cognition and Neuroscience*. **35**(1), pp.17–42.
- King, J. and Just, M. 1991. Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language*. **30**, pp.580–602.
- King, J. and Kutas, M. 1995. Who did what and when? Using word- and clause-level ERPs to monitor working memory usage in reading. *Journal of Cognitive Neuroscience*. (7), pp.376–395.
- Kluender, R. 1991. *Cognitive constraints on variables in syntax*. Ph.D. thesis, University of California at San Diego.
- Kluender, R. 1998. On the distinction between strong and weak islands: A processing

- perspective. In: P. Culicover and L. McNally, eds. *Syntax and Semantics: The Limits of Syntax*. New York: Academic Press, pp.241–280.
- Kluender, R. and Kutas, M. 1993a. Bridging the gap: Evidence from ERPs on the processing of unbounded dependencies. *Journal of Cognitive Neuroscience*. **5**(2), pp.196–214.
- Kluender, R. and Kutas, M. 1993b. Subjacency as a processing phenomenon. *Language and Cognitive Processes*. **8**(4), pp.573–633.
- Kroch, A. 1981. On the role of resumptive pronouns in amnestying island constraint violations. In: R. A. Hendrick, C. S. Masek and M. F. Miller, eds. *Papers from the seventeenth regional meeting*. Chicago: Chicago Linguistic Society, pp.125–135.
- Kush, D., Lidz, J. and Phillips, C. 2015. Relation-sensitive retrieval: Evidence from bound variable pronouns. *Journal of Memory and Language*. **82**, pp.18–40.
- Kush, D., Lohndal, T. and Sprouse, J. 2015. *Experimental syntax and the cross-linguistic variation of island effects in Norwegian and Swedish*. [Poster]. Exhibited at Linguistic Society of America (LSA).
- Kush, D., Lohndal, T. and Sprouse, J. 2017. Investigating variation in island effects: A case study of Norwegian. *Natural Language and Linguistic Theory*. **36**(3), pp.743–779.
- Lavine, J.E. 2003. Resumption in Slavic: phases, cyclicity, and case. *Formal Approaches to Slavic Linguistics*. **11**, pp.355–372.
- Lewis, R.L. and Vasishth, S. 2005. An activation-based model of sentence processing as skilled memory retrieval. *Cognitive science*. **29**(3), pp.375–419.
- Lewis, S. and Phillips, C. 2014. Aligning grammatical theories and language processing models. *Journal of Psycholinguistic Research*. **44**(1), pp.27–46.
- Maamouri, M. 1998. *Language education and human development: Arabic diglossia and its impact on the quality of education in the Arab region*. Washington, D. C.: The World Bank.
- Maling, J. and Zaenen, A. 1982. A phrase structure account of Scandinavian extraction phenomena. In: P. Jacobson and G. rey K. Pullum, eds. *The Nature of Syntactic Representation*. Dordrecht: Reidel, pp.229–282.
- Martorell, J. 2018. Merging generative linguistics and psycholinguistics. *Frontiers in*

- Psychology*. **9**(NOV), pp.1–5.
- Mathôt, S., Schreij, D. and Theeuwes, J. 2012. OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*. **44**(2), pp.314–324.
- McCloskey, J. 2017. New thoughts on old questions—resumption in Irish. In: *Asking the right questions: Essays in honor of Sandra Chung* [Online]., pp.81–102. Available from: <https://escholarship.org/content/qt8255v8sc/qt8255v8sc.pdf#page=89>.
- McCloskey, J. 2002. Resumption, successive cyclicity, and the locality of operations. In: S. D. Epstein and D. Seely, eds. *Derivation and Explanation in the Minimalist Program*. Oxford: Blackwell, pp.84–226.
- McCloskey, J. 2006. Resumption. In: M. Everaert and H. R. Van, eds. *Blackwell companion to syntax*,. Oxford: Blackwell, pp.94–115.
- McCloskey, J. 1990. Resumptive pronouns, A-binding and levels of representation in Irish. In: R. Hendrick, ed. *The Syntax of the Modern Celtic Languages*. New York and San Diego: Academic Press, pp.199–248.
- McCloskey, J. 1979. *Transformational syntax and model theoretic semantics: a case-study in Modern Irish*. Dordrecht: Reidel.
- Meltzer-Asscher, A. 2018. *Resumption, what is it good for? Absolutely nothing?* [Online]. [PowerPoint presentation]. Available from: <https://osf.io/7s3uz/>.
- Meltzer-Asscher, A., Fadlon, J., Goldstein, K. and Holan, A. 2015. Direct object resumption in Hebrew: How modality of presentation and relative clause position affect acceptability. *Lingua*. **166**, pp.65–79.
- Merchant, J. 2004. Resumptivity and non-movement. *Studies in Greek Linguistics*. **24**, pp.471–481.
- Michel, D. 2014. *Individual cognitive measures and working memory accounts of syntactic island phenomena*. Ph.D. thesis, University of California , San Diego.
- Mohammad, M. 2000. *Word order, agreement and pronominalization in Standard and Palestinian Arabic*. Amsterdam: John Benjamins.
- Myers, J. 2009. Syntactic judgment experiments. *Linguistics and Language Compass*. **3**(1), pp.406–423.
- Omaki, A. and Schulz, B. 2011. Filler-gap dependencies and island constraints in

- second-language sentence processing. *Studies in Second Language Acquisition*. **33**(4), pp.563–588.
- Pañeda, C., Lago, S., Vares, E., Veríssimo, J. and Felser, C. 2020. Island effects in Spanish comprehension. *Glossa: a journal of general linguistics*. **5**(1), p.21.
- Patterson, C. 2013. Discourse coherence and pronoun resolution. *Discours*. **12**(12).
- Perpiñán, S. 2020. Wh-Movement, islands, and resumption in L1 and L2 Spanish: Is (Un)grammaticality the relevant property? *Frontiers in Psychology*. **11**(March), pp.1–13.
- Pesetsky, D. 1987. Wh-in-situ: Movement and unselective binding. *The representation of (in) definiteness*. **98**, pp.98–129.
- Phillips, C. 2013. On the nature of island constraints. I: Language processing and reductionist accounts. In: J. Sprouse and N. Hornstein, eds. *Experimental syntax and island effects*. Cambridge: Cambridge University Press, pp.64–108.
- Phillips, C. 2009. Should we impeach armchair linguists? *Japanese/Korean Linguistics*. **17**, pp.1–16.
- Phillips, C., Gaston, P., Huang, N. and Muller, H. 2019. Theories all the way down: remarks on "theoretical" and "experimental" linguistics." In: G. Goodall, ed. *Cambridge Handbook of Experimental Linguistics*.
- Phillips, C. and Wagers, M.W. 2007. Relating structure and time in linguistics and psycholinguistics. In: G. Gaskell, ed. *Oxford Handbook of Psycholinguistics*., pp.739–756.
- Polinsky, M. 2011. Reanalysis in adult heritage language: New evidence in support of attrition. *Studies in Second Language Acquisition*. **33**(2), pp.305–328.
- Polinsky, M., Lauren Eby Clemens, Adam M. Morgan, Ming Xiang and David Heestand. 2013. Resumption in English. In: Jon Sprouse and Norbert Hornstein, eds. *Experimental syntax and island effects*., Cambridge University Press, pp.341–360.
- Prince, E.F. 1990. Syntax and discourse : A look at resumptive pronouns. *Berkeley Linguistics Society*. **16**, pp.482–497.
- Reitbauer, M. 2013. *Resumption, movement effects, and a matching analysis of A'-Dependencies*. Ph.D. thesis, uniwiien.
- Rizzi, L. 2004. Locality and left periphery. *tructures and beyond: The cartography of*

syntactic structures. **3**, pp.223–251.

Rizzi, L. 2001. On the position “Int (errogative)” in the left periphery of the clause. *In*: G. Cinque and G. Salvi, eds. *Current studies in Italian syntax: Essays offered to Lorenzo Renzi*. Amsterdam: Elsevier, pp.287–296.

Rizzi, L. 1990. *Relativized minimality*. The MIT Press.

Ross, J. 1986. *Infinite syntax*. Norwood: NJ: Ablex.

Ross, J.R. 1967. *Constraints on variables in syntax*. Ph.D. thesis, MIT.

Rouveret, A. 2011. Some issues in the theory of resumption. A perspective on early and recent research. *In*: A. Rouveret, ed. *Resumptive pronouns at the interfaces*. Amsterdam, the Netherlands: John Benjamins, pp.1–62.

Saeed, J. 1984. *The syntax of focus and topic in Somali*. Hamburg: Buske.

Salzmann, M. 2009. When movement and base-generation compete: The definition of the reference set and parameterized preferences for elementary operations. *Linguistics in the Netherlands*. AVT Publications. **26**, pp.64–77.

Schütze, C.T. 1996. *The empirical base of linguistics*. University of Chicago Press.

Sells, P. 1987. Binding resumptive pronouns. *Linguistics and Philosophy*. **10**(3), pp.261–298.

Sells, P. 1984. *Syntax and semantics of resumptive pronouns*. Ph.D. thesis, University of Massachusetts at Amherst.

Shapiro, L.P. 2000. Some recent investigations of gap-filling in normal listeners: Implications for normal and disordered language processing. *In*: L. Y. Grodzinsky, Shapiro and D. Swinney, eds. *Language and the brain: Representation and processing*. San Diego: Academic Press, pp.357–376.

Sharvit, Y. 1999. Resumptive pronouns in relative clauses. *Natural Language & Linguistic Theory*. **17**(3), pp.587–612.

Shlonsky, U. 1992. Resumptive pronouns as a last resort. *Linguistic Inquiry*. **23**(3), pp.443–468.

Soltan, U. 2007. *On formal feature licensing in minimalism: Aspects of Standard Arabic morphosyntax*. Ph.D. thesis, University of Maryland.

Sorace, A. and Keller, F. 2005. Gradience in linguistic data. *Lingua*. **115**(11),

pp.1497–1524.

- Sprouse, J. 2007. *A program for experimental syntax: finding the relationship between acceptability and grammatical knowledge*. Ph.D. thesis, University of Maryland.
- Sprouse, J., Caponigro, I., Greco, C. and Cecchetto, C. 2016. Experimental syntax and the variation of island effects in English and Italian. *Natural Language and Linguistic Theory*. **34**(1), pp.307–344.
- Sprouse, J. and Hornstein, N. (eds.). 2013. *Experimental syntax and island effects*. Cambridge: Cambridge University Press.
- Sprouse, J. and Schütze, C.T. 2017. Grammar and the use of data. In: B. Aarts, J. Bowie and G. Popova, eds. *Oxford Handbook of English Grammar*. Oxford: Oxford University Press.
- Sprouse, J., Wagers, M. and Phillips, C. 2012. A test of the relation between working memory and syntactic island effects. *Language*. **88**, pp.82–123.
- Sprouse, J., Wagers, M.W. and Phillips, C. 2013. Deriving competing predictions from grammatical approaches and reductionist approaches to island effects. In: J. Sprouse and N. Hornstein, eds. *Experimental Syntax and Island Effects*. Cambridge: Cambridge University Press, pp.21–41.
- Sterian, L. 2016. *Obligatory And optional resumption; Case studies in the syntax of Iraqi Arabic and Romanian*. Ph.D. thesis, University of Edinburgh.
- Stewart, A.J., Holler, J. and Kidd, E. 2007. Shallow processing of ambiguous pronouns: Evidence for delay. *The Quarterly Journal of Experimental Psychology*. **60**(12), pp.1680–1696.
- Stoet, G. 2017. PsyToolkit: A novel web-based method for running online questionnaires and reaction-time experiments. *Teaching of Psychology*. **44**(1), pp.24–31.
- Stoet, G. 2010. PsyToolkit - A software package for programming psychological experiments using Linux. *Behavior Research Methods*. **42**(4), pp.1096–1104.
- Stowe, L.A. 1986. Parsing WH-constructions: Evidence for on-line gap location. *Language and Cognitive Processes*. **1**(3), pp.227–245.
- Sulaiman, M. 2016. *The syntax of wh-questions in syrian arabic*. Ph.D. thesis, Newcastle University.

- De Swart, H. 1992. Intervention effects, monotonicity and scope. *Semantics and Linguistic Theory*. **2**, pp.387–406.
- Szabolcsi, A. and Lohndal, T. 2017. Strong vs. weak islands. *In*: M. Everaert and H. van Riemsdijk, eds. *The Wiley Blackwell Companion to Syntax*., pp.1–51.
- Szabolcsi, A. and Zwarts, F. 1993. Weak islands and an algebraic semantics for scope taking. *Natural Language Semantics*. **1**(3), pp.235–284.
- Traxler, M.J. and Pickering, M.J. 1996. Plausibility and the processing of unbounded dependencies: an eye-tracking study. *Journal of Memory and Language*. **35**(3), pp.454–475.
- Troyer, M., Hofmeister, P. and Kutas, M. 2016. Elaboration over a discourse facilitates retrieval in sentence processing. *Frontiers in Psychology*. **7**(MAR), pp.1–9.
- Truswell, R. 2007. Extraction from adjuncts and the structure of events. *Lingua*. **117**(8), pp.1355–1377.
- Tsimpli, I.M. 2003. Interrogatives in the Greek/English interlanguage: A minimalist account. *Selected papers on theoretical and applied linguistics*. **15**, pp.214–225.
- Tsimpli, I.M. 1999. Null operators, clitics and identification: A comparison between Greek and English. *In*: R. Alexiadou, G. Horrocks and M. Stavrou, eds. *Studies in Greek Syntax*. Dordrecht: Kluwer, pp.241–262.
- Tucker, Matthew A, Idrissi, A., Sprouse, J. and Almeida, D. 2019. Resumption ameliorates different islands differentially: Acceptability data from Modern Standard Arabic. *In: Perspectives on Arabic Linguistics XXX: Papers from the annual symposia on Arabic Linguistics, Stony Brook, New York, 2016 and Norman, Oklahoma, 2017*. John Benjamins Publishing Company.
- Wagers, M.W. and Phillips, C. 2014. Going the distance: Memory and control processes in active dependency construction. *Quarterly Journal of Experimental Psychology*. **67**(7), pp.1274–1304.
- Wagers, M.W. and Phillips, C. 2009. Multiple dependencies and the role of the grammar in real-time comprehension. *Journal of Linguistics*. **45**(2), pp.395–433.
- Wanner, E. and Maratsos, M. 1978. An ATN approach to comprehension. *In*: Morris Halle, Joan Bresnan and George Miller, eds. *Linguistic theory and psychological reality*. Cambridge, MA: MIT Press, pp.119–61.

- Wood, S.N. 2011. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society, Series B.* **73**(1), pp.3–36.
- Woods, D.L., Kishiyama, M.M., Yund, E.W., Herron, T.J., Edwards, B., Poliva, O. and Reed, B. 2011. Improving digit span assessment of short-term verbal memory. *Journal of clinical and experimental neuropsychology.* **33**, pp.101–111.
- Yoshida, M., Kazanina, N., Pablos, L. and Sturt, P. 2014. On the origin of islands. *Language, Cognition and Neuroscience.* **29**(7), pp.761–770.

APPENDIX A Experimental Sentences

Sentences included in Study 1

Non-Crossed Island:

A. Bare WH-questions:

1. ?ayš amal zar-at-/uh lamman kan-at fi ar-ryaḍ?

What Amal visited-3SF-/=it when was-3SF in riyadh?

What did you visit when you were in Riyadh?

2. ?ayš raḥ taḥt-in-/uh mḥammad law nejeḥ fi

what will give-2SF-/=it Muhammad if passes.3SM in

al-ekhtebar?

the-exam?

What would you give Muhammad if he passed the exam?

3. ?ayš kan mḥammad yaḥtaj-/uh qablma

what was.3SM Muhammad want.3SM-/=it before

ysafer?

travelled.3SM?

What did Muhammad need before he travelled?

4. ?ayš raḥ amal tulqi-/uh law ʃarak-at fii

what will Amal recite.3SF-/=it if participated-3SF in

al-musabaq al-ʔdabyyah?

the-competition the-literature?

What would Amal recite if she participated in the literature competition?

5. ?ayš khaled kasar-/uh lamman tḍarab maḥ

what Khaled broke.3SM-/=it when fought.3SM with
aṣḥab-uh?

frinds-his?

What did Khaled break when he fought with his friends?

6. ḡayš raḥ afnan talbas-/=uh law raḥ-at

what will Afnan wear.3SF-/=it if went-3SF
az-zawaj?

the-wedding-party?

What will Afnan wear if she attended the wedding party?

B. Cleft WH-questions:

1. ḡayš al-fustan illi lebes-ti-/=h lamman ru-ḥti

What the-dress that wore-2SF-/=it when attended-2SF
al-ḥaflah?

the-party?

What is the dress that you wore when you attended the party?

2. ḡayš al-kukh illi fahad əstʔjar-/=uh lamman kan

what the-cottage that Fahad rented.3SM-/=it when was.3SM
fii turkia?

in Turkey?

What is the cottage that Fahad rented when he was in Turkey?

3. ḡayš al-maṭʕam illi maha zar-at-/=uh lamman kan-at

what the-resturant that Maha visited-3SF-/=it when was-3SF
fii Jeddah?

in jeddah?

What is the restaurant that Muna visited when she was in Jeddah?

4. ?ayš al-musabaqah illi al-laʿeb kheser-/=ha
what the-competition that the-player loose.3SM-/=it
bi-sabab al-?šaba?
because the-injury?

What is the competition that the-player loose because of the injury?

5. ?ayš as-saʿah illi khaled lebes-/=ha lamman
what the-watch that Khaled wore.3SM-/=it when
raḥ al-ḥafl?
went.3SM the-party?

What is the watch that Khaled wore when he attended the party?

6. ?ayš at-taqreer illi aš-šaḥīfah naʿjar-at-/=uh lamman
what the-report that the-newspaper published-3SF-/=it when
estaqal al-mudeer?
retired.3SM the-administrator?

What is the report that the newspaper published when the administrator retired?

C. Relative Clauses

1. ḥabai-t aš-šuqah illi ?st?jar-ti-/=ha lamman
liked-1S the-flat that rented-2SF -/=it when
ruḥ-ti Turkia
went-2SF Tutkey

I liked the flat that you rented when you went to Turkey

2. ḥaḍar-t al-musabaqah illi amal nḍam-at-/=ha
attended-1S the-competition that Amal organized-3F-/=it
baḍma rejeḥ-at men America
after come.back-3SF from America
I attended the competition that Amal organized after she came back from America

3. ḥaba-it al-esbaqatti illi amal tswwt-/=ha lamman
liked-1S the-spagetti that Amal prepared-3sF-/=it when
zur-na-ha fii jeddah
visited-1PL-her in Jeddah.
I liked the espagetti that Amal prepared when we visited her in jeddah

4. əstakhdam-t ad-darrajah illi saleh iftra-/=ha lamman
used-1S the-bicycle that Saleh bought.3SM-/=it when
kan fii Abha.
was.3SM in Abha
I rode the bicycle that Saleh bought when he was in Jeddah

5. Maḥammad kasar as-saḥah illi huda jab-at-/=ha
Muhammad broke.3SM the-watch that Huda broght-3SF-/=it
lamman ḥbu-yah taqaḥad
when father-my retired.3SM
Muhammad broke he watch that Huda brought when my father retired.

6. əstakhdam-t al-computer illi maha barmaj-at-/=uh baḍma
used-1S the-computer that Maha programmed-3SF-/=it after
tḥaḥal

was-down

I used the computer that Maha programmed when it was down

Crossed Non-Island:

A. Bare WH-questions:

1. ?ayš tetwaqƣ-een ?nn mħmmad raħ yebeeƣ-/uh?

what think-2SF that Muhammad will sell.3SM-/it?

What do you think that Muhammad will sell?

2. ?ayš yabġa khaled yarsum-/uh?

what want.3SM Khaled draw.3SM-/it

What does Khaled want to draw?

3. ?ayš tetwaqaƣ-een inn maha raħ tenaqif-/uh?

what think-.2SF that Maha will discuss.3SF-/it

What do you think that Maha will discuss?

4. ?ayš qal aħmad inn al-baladyyah raħ tebni-
/h?

what said.3SM Ahmad that the-council will build.3SM-/it?

What did Ahmad say that the council will build?

5. ?ayš tetwaqaƣ=een inn ar-rassam bi-yarsum-/uh?

what think-2SF that the-artist will-draw.3SM-/it?

What do you think that the artist will draw?

6. ?ayš maha qal-at inn-ha raħ tejeeb-/uh?

what Maha said-3SF that-she will bring.3SF-/it?

What did Maha say that she will bring?

B. Cleft WH-questions:

1. ʔayš as-syyarah illi qul-ti ʔnn ahamad sadam-/=ha?

What the-car that said-2SF that Ahmad hit.3SM-/=it?

What is the car that you said that Muhammad hit?

2. ʔayš as-syyara illi al-hurras əṣṭraf-u inn
waḥed

what the-car that the-guards admitted-3PLM that one

men-hum kasar-/=ha?

of-them broke.3SM-/=it?

What is the car that the guards admitted that one of them has broken?

3. ʔayš al-qaryah illi qul-ti l-e inn ʔumm-ek

what the-village that said-2SF for-me that mother-your

zar-at-/=ha?

visited-3SF-/=it?

what is the village that you told me that your mother visited?

4. ʔayš al-mustaffa illi amal qal-at inn ʔkhu-ha

what the-hospital that Amal said-3SF that brother-her

yudeer-/=uh?

manage.3SM-/=it?

What is the hospital that Amal said that your brother manage?

5. ʔayš al-jareemah illi qul-ti inn al-mudarres

what the-crime that said-2SF that the-teacher

ertakab-/=ha?

commit.3SM-/=it?

What is the crime that you said that the teacher commits?

6. ?ayš al-mabna illi qul-ti inn saleh raḥ

What the-building that said-2SF that saleh will

grammem-/=uh?

refurbish.3SM-/=it?

What is the building that you said that Saleh will refurbish?

C. Relative Clauses

1. šef-t as-saḡah illi qul-ti l-ee ?nn saleḥ

saw-1s the-watch that told-2SF for-me that Saleh

?štra--/ha

bought.3SM-/=it

I saw the watch that you told me that Saleh bought -/it

2. šef-t as-sayyara illi khaled qal inn mḥammad

saw-1S the-car that Khaled said.3SM that muhammad

šadam-/=ha

hit.3SM-/=it

I saw the car that Khaled said that Muhammad hit

3. zur-t al-qariah illi qul-ti inn ?mm-ek

visit-1S the-village that said-2SF that mother-your

tḥubb-/=ha

like.3SF-/=it.

I visited the village that you said that your mother loves.

4. ?ʕjab-ni al-qaʕr illi qul-ti l-e inn ?khu-k
liked-1S the-palace that told-2SF for-me that brother-your
ʕammam-/uh
designed.3SM-/it.

I liked the palace that you told me that your brother designed

5. ʕabbai-t al-film illi muna qal-at l-e inn-ha
loved-1S the-movie that Muna told-3SF for-me that-she
tabaʔ-at-/uh
watched-3SF-/it.

I liked the movie that Muna told me that she watched.

6. ʕaddaq-t bi-al-fustan illi huda qal-at inn-ha
donated-1S by-the-dress that huda said-3SF that-she
ʕamm-at-/uh
designed-3SF-/it

I donated the dress that Huda said that she designed.

Crossed Adjunct-Island:

A. Bare WH-questions:

1. ?ayš muna ʕaḍar-at al-muḥaḍarah qablma taqrʔ-/uh?
what Muna attended-3SF the-lecture before read.3SF -/it?
What did Mona attended the lecture before she read?
2. ?ayš kammal-ti ʕal al-ʔsʔelah bidunma
what finished-2SF answering the-questions without
tlaḥz-een-/uh?

notice-2SF-/=it?

What did you finish answering the questions before you noticed?

3. ?ayš aš-šurṭah mesik-uu khaled lamman
what the-police arrested-3PLM Khaled when
əstraq-/=uh?
stole.3SM-/=it?

What did the police arrest Khaled when he stole?

4. ?ayš al-mudeer faṣal al-muḏaf lamman
what the-administrator sacked.3SM the-employee when
ahmal-/=uh?
dismissed.3SM-/=it?

What did the administrator sack the employer when he dismissed?

5. ?ayš ahmad ekhtar al-muhandes baḏma ṣammem-/=uh?
what Ahmad chose.3SM the-engineer after designed.3SM-/=it?
What did Ahmad choose the engineer after he designed?

6. ?ayš maha zeḡel-et lamman amal kasar-at-/=uh?
what Maha felt-angry-3F when Amal broke.3F-/=it?
What did Maha become angry when Amal broke?

B. Cleft WH-questions:

1. ?ayš al-maqal illi bent-k faz-at bi-al-musabaqah
What the-article that daughter-your won-3SF by-the-competition
lamman qar?-at-/=uh?
when read-3SF -/=it?

What is the article that your daughter won the competition when she read?

2. ?ayš al-lawḥah illi aš-šurṭah mesek-uu khaled
what the-painting that the-police arrested-3PLM Khaled
lamman əstrag-/=ha?
when stole.3SM-/=it?

What is the painting that the police arrested Khaled when he stole?

3. ?ayš al-baranamej illi huda zeʕel-t ʕlašan
what the-show that Huda felt-angry.3SF because
bent-ha tabaʕ-at-/=uh?
daughter-her watched-3SF-/it?

What is the show that Huda became angry because her daughter watched?

4. ?ayš al-qeṭʕah illi al-jehaz khereb lamman
aḥmad
what the-piece that the-machine damaged.3SM when Ahmad
kasar-/=ha?
broke.3SM-/=it?

What is the piece that the-machine damaged when Ahmad broke?

5. ?ayš aš-šuruṭ illi rah yefraj-uu ʕan ahmad law
what the-conditions that will release-3PLM about Ahmad if
əjtaz-/=ha?
met.3SM-/=it?

What are the conditions that they will release Ahmad if he met?

6. ?ayš al-madinah illi inbasaṭ-i lamman

what the-city that felt-happy.2SF when

zur-ti-/=ha?

visited-2SF-/=it?

What is the city that you had fun when you visited?

C. Relative Clauses

1. kabbai-t al-ḥaleeb illi khaled meriḍ baʿd-ma

spilled-1S the-milk that Khaled felt-sick.3SM after

šerib -/=uh

drank.3SM-/=it

I spilled the milk that Khaled became sick after he drank

2. šef-t al-jeſr illi mḥammad safar qablma

saw.1S the-bridge that Mhammad travelled.3SM before

yhdem-un-/=h

destroy-3LPM-/=it

I saw the bridge that Muhammad travelled before they destroy.

3. əštrai-t al-lawḥah illi aṭ-ṭullab inbahar-uu

bought-1S the-painting that the-students astonished-3PLM

lamman šaf-au-/=ha

when saw-3PLM-/=it

I bought the painting that the students were astonished when they saw

4. ʔkal-t al-ḥala illi mḥammad inbašaṭ lamman

ate-1S the-dessert that Mhammad felt-happy.3M when

jeb-t-/=uh

brought-1S-/=it .

I ate the dessert that Muhammad became happy when I brought-/it.

5. əštrait al-ketab illi šakħşyyat muna teħassan-at
bought-1S the-book that personality Muna improved-3SF
lamman qraʔ-t-/uh
when read-3SF-/it

I bought the book that Muna's personality improved when she read.

6. reħ-t al-estraħah illi saleħ raħ yeʕzem ašħabuh
went-1S the-lounge that Saleh will invite.3SM friends-
his
law estʔjar-/ha
if rent.3SM-/it

I went to the lounge that Saleh would invite his friends if he could rent.

Crossed RC-Island:

A. Bare WH-questions:

1. ʔayš mħmmad ʕazam al-fannan illi rasam-/uh?
what Muhammad invited.3SM the-artist who drew.3SM -/it?
What did Muhammad invite the artist who drew?

2. ʔayš saleħ yeʕref ar-rijjal illi əštra-/h?
what Saleh know.3SM the-man that bought.3SM-/it?
What does Saleh know the man who bought?

3. ʔayš khaled qabal al-mukhrej illi akhraj-/uh?
what Khaled met.3SM the-director that direct.3SM-/it?
What did Khaled meet the director who directed?

4. ?ayš barak-ti l-bent illi ektašaf-at-/uh?
what congratulate.2SF the-girl that invented-3SF-/=it?
What did you congratulate the girl who invented?

5. ?ayš ali šakar al-muhandis illi šammam-/uh?
what Ali thanked.3SM the-engineer who designed.3SM-/=it
What did Ali thank the engineer who designed?

6. ?ayš aš-šurṭah mesik-u al-ḥarami illi
what the-police arrested-3PLM the-thief who
əstraq-/uh?
stole.3SM-/=it?
What did the police arrest the thief who stole?

B. Cleft WH-questions:

1. ?ayš al-jihaz illi al-mudeer karram
What the device that the-administrator honoured.3SM
al-mozaḥf illi šanaḥ -/uh?
the-employer who made.3SM-/=it?
What is the device that the administrator honoured the employer who made?

2. ?ayš al-kitab illi saleḥ yeḥref ar-rijjal illi
what the-book that Saleh know.3SM the-man who
?llaf-/uh?
wrote.3SM-/=it?
What is the book that Saleh knows the man who wrote?

3. ?ayš al-lawḥah illi fahad ʕazam ar-rejjaḷ illi
what the-painting that Fahad invited.3SM the-man who
rasam-/=ha?

draw.3SM-/=it?

What is the painting that Fahad invited the man who drew?

4. ?ayš al-jehaz illi almudeer karram aṭ-ṭaleb
what the-machine that the-administrator honored.3SM the-
student

illi šammam-/=uh?

who designed.3SM-/=it?

What is the machine that the administrator honored the student who invented?

5. ?ayš as-sayyara illi al-mujrem qatal al-walad
what the-car that the-criminal killed.3SM the-boy
illi kan ysuq-/=ha?

who was.3SM driving.3SM-/=it?

What is the car that the criminal killed the boy who was driving?

6. ?ayš al-masraḥyyah illi snaa teʕref al-kateb-ah
what the-play that Snaa know.3SF the-athour-F
illi allf-at-/=ha?

who wrote-3SF-/=it?

What is the play that Sanaa knows the authour who wrote?

C. Relative Clauses

1. ?kal-t al-ḥala illi muna teʕref al-bent illi

ate-1S the-dessert that Muna knew.3SF the-girl that
saww-at-/=uh
made-3SF-/=it

I eat the dessert that Mona knows the girl who made

2. mḥammad yudrus fii al-madrasah illi saleḥ yeʕref
Mhammad study.3SM in the-school that saleh know.3SM
ar-jjal illi yamlek-/=ha
the-man who owned.3SM-/=it

Muhammad is studying at the school that Saleh knows the man who own

3. ḥabbai-t al-fellah illi qabl-na ar-rejjal illi
liked-1S the-vella that met-we the-man who
ṣammam-/=ha
designed.3SM-/=it

I liked the vella that we met the man who designed

4. šreb-t al-ʕaṣeer illi ahmad tsamma baʕdma
drank-1S the-juice that Ahmad poisoned.3M after
šerb-/=uh
drank.3SM-/=it

I drank the juice that Ahmad was poisoned after he drank

5. dawwar-t ʕla al-ketab illi neʕref at-ṭaleb
lookerd-1S for the-book that know.1PL the-student
illi ʔstaʕar-/=uh
who borrowed.3SM-/=it

I looked for the book that we know the student who borrowed.

6. ḡhub al-ḡuḡniah illi muna teḡref al-muṡreb illi
liked-1S the-song that Muna know.3SF the-singer who
laḡan-/=ha
wrote.3SM-/=it

I like the song that Muna knows the singer who wrote

Sentences included in Study 2

Non-Crossed Island:

A. Bare WH-questions:

1. ?ayš amal zar-at-/uh lamman kan-at fi
What amal visit-3SF-/it when was-3SF in
ar-ryaḍ?

the-Riyadh?

What did you visit when you were in Riyadh?

2. ?ayš raḥ taṣṭ-in-/uh mahmmad law nejeḥ fi
what will give-2SF-/=it Muhammad if passes.3SM in
al-ekhtebar?

the-exam?

What would you give Muhammad if he passed the exam?

3. ?ayš kan mḥammad yaḥtaj-/uh qablma ysafer
what was Muhammad need.3SM-/=it before travel.3SM
jeddah?

Jeddah?

What did Muhammad need before he travelled to Jeddah?

4. ?ayš raḥ amal tulqi-/=h law šarak-at fii
what will Amal recite.3SF-/=it if participated-3SF in
al-musabaqah al-ʔdabyyah?

the-competition the-literature?

What would Amal recite if she participated in the literature competition?

5. ?ayš khaled kasar-/uh lamman tḍarab maṣ

what Khaled broke.3SM-/=it when fought.3SM with
aṣḥab-uh?

frinds-his?

What did Khaled break when he fought with his friends?

6. ʔayš raḥ afnan talbas-/=uh law raḥ-at
what will Afnan wear.3SF-/=it if went-3SF
az-zawaj?

the-wedding-party?

What will Afnan wear if she attends the wedding party?

B. Relative Clauses

1. ḥabai-t al-šūqah illi ʔstʔjar-ti-/=ha lamman
liked-1S the-flat that rented-2SF -/=it when
ruḥ-ti turkia
went-2SF Tutkey

I liked the flat that you rented when you went to Turkey

2. ḥaḍar-t al-musabaqah illi amal nḏam-at-/=ha
attended-1S the-competition that amal organized-3F-/=it
baʕdma rejeʕ-at men america
after come.back-3SF from America

I attended the competition that Amal organized after she came back from America

3. ḥabai-t al-esbaqatti illi amal sww-at-/=ha lamman
liked-1S the-spaghetti that Amal prepared-3SF-/=it when
zur-na-ha fii jeddah

visited-1PL-her in Jeddah.

I liked the spaghetti that Amal prepared when we visited her in Jeddah.

4. əstakhdam-t ad-darrajah illi saleh ištra-/=ha lamman
used-1S the-bicycle that Saleh bought.3SM-/=it when
kan fii abha.
was.3SM in Abha

I rode the bicycle that Saleh bought when he was in Abha.

5. Maḥammad kasar as-saḥah illi huda jab-at-/=ha
Muhammad broke.3SM the-watch that Huda brought-3SF-/=it
lamman ʔbu-yah taqaḥad
when father-my retired.3SM

Muhammad broke the watch that Huda brought when my father retired.

6. əstakhdam-t al-computer illi maha barmaj-at-/=uh baḥdma
used-1S the-computer that Maha programmed-3SF-/=it after
tḥaṭal ʔms
was-down yesterday

I used the computer that Maha programmed when it was down yesterday

Crossed Non-Island:

A. Bare WH-questions:

1. ʔayš tetwaḥḥ-een ʔnn mḥmmad raḥ yebeeḥ-/=h fii
what think-2SF that Muhammad will sell.3SM-/=it in
mḥall-uh al-jdeed?
shop-his the-new?

What do you think that Muhammad will sell in his new shop?

2. ?ayš yabġa khaled yarsum-/uh řla jedar al-fařl?
what want.3SM Khaled draw.3SM-/it on wall the-classroom

What does Khaled want to draw on the classroom wall?

3. ?ayš tetwaqař-eeen inn maha raġ tenaqif-/uh fi ejtemař
what think-.2SF that Maha will discuss.3SF-/=it in
meeting al-edarah?

the-administration?

What do you think that Maha will discuss in the administration meeting?

4. ?ayš qal aġmad inn al-baladyyah raġ tebni-/=h fii
what said.3SM Ahmad that the-council will build.3SM-/=it in
wařt al-madiinah?

center the-city?

What did Ahmad say that the council will build in city center?

5. ?ayš tetwaqař-eeen inn ar-rassam bi-yarsum-/uh řla
what think-2SF that the-artist will-draw.3SM-/=it on
al-jedar al-khareji?

the-wall the-outer?

What do you think that the artist will draw on the outer wall?

6. ?ayš tetwaqař-un inn maha bi-tquul-/uh bařd
what think-3SF that Maha will-say.3SF-/=it after
ġflah az-zawaj?

party the-wedding?

What do you think that Maha will say after the wedding party?

B. Relative Clauses

1. šef-t as-saḥah illi qulti l-e ʔnn saleh
saw-1s the-watch that told-2SF for-me that Saleh
ʔštra-/=ha li-haflat takharrij amal
bought.3SM-/=it for-party graduation Amal

I saw the watch that you told me that Saleh bought -/it for Amal's graduation party

2. šef-t as-sayyara illi khaled qal inn mḥammad
saw-1S the-car that Khaled said.3SM that Muhammad
šadam-/=ha fii ṭreeq al-ḡabah
hit.3SM-/=it in way the-forest.

I saw the car that Khaled said that Muhammad hit -/it in his way to the forest

3. zur-t al-qariah illi qul-ti inn ʔmmek
visit-1S the-village that said-2SF that mother-your
tzur-/=ha fi kul šaif
visits.3SF-/=it in every summer

I visited the village that you said that your mother visits -/it every summer.

4. habbai-t al-qaṣr illi qulti inn ʔkhuk
liked-1S the-palace that saids-2SF that brother-your
šammam-/=uh li-mudeer šarekat al-esment
designed.3SM-/=it for-maneger company the-cement.

I liked the palace that you told me that your brother designed -/it for the cement company manager

5. ḥabbai-t al-film illi muna qal-at inn-ha

loved-1S the-movie that muna said-3SF that-she
tabaʔ-at-/uh mʕa ʕaḥbat-ha ʔms
watched-3SF-/it with frinds-her yesterday.

I liked the movie that Muna told me that she watched -/it with her friends

6. ʕaɖdaq-t bi-alfustan illi huda qal-at inn-ha
donated-1S by-the-dress that Huda said-3SF that-she
ʕammat-/uh li-haflat zawaj ukht-ha
designed-3SF-/it for-party wedding sister-her

I denoted the dress that Huda said that she designed -/it for her sister's
wedding party

Crossed Adjunct-Island:

A. Bare WH-questions:

1. ʔayš muna ḥaɖar-at al-muḥaɖarah qablma taqrʔ-/h
what Muna attended-3SF the-lecture before read.3SF -/it
fii maktabat al-jameʕah?
in library the-university?

What did Mona attended the lecture before she read in the university library?

2. ʔayš kammal-ti ḥal al-ʔsʔelah bidunma
what finished-2SF answering the-questions without
tlaḥz-eeen-/uh fii nehayat al-ekhtebar?
notice-2SF-/it at end the-exam?

What did you finish answering the questions before you noticed at the end of
the exam?

3. ʔayš aš-šurḥah mesik-uu al-ḥarami lamman

what the-police arrested-3PLM the-thief when
əstrɑq-/=uh min baitu-kum al-jdeed?
stole.3SM-/=it from house-your the-new?

What did the police arrest the thief when he stole from your new house?

4. ʔayš al-mudeer faṣal al-muṣaf lamman
what the-administrator sacked.3SM the-employee when
ahmal-/=uh fii akher šahrain?
dismissed.3SM-/=it in last month.DU?

What did the administrator sack the employee when he dismissed in the last two months?

5. ʔayš ahmad ekhtar al-muhandes baʕdma šammem-/=uh
what Ahmad chose.3SM the-engineer after designed.3SM-
/=it
fii mul al-ʕrab?
in mall the-Arab?

What did Ahmad choose the engineer after he designed in Arab Mall?

6. ʔayš maha zeʕel-et lamman amal kasar-at-/=uh
what maha felt-angry-3F when Amal broke.3F-/=it
fii ḥaflat ʔms?
in party yesterday?

What did Maha become angry when Amal broke in yesterday's party?

B. Relative Clauses

1. kabbai-t al-ḥaleeb illi khaled meriḍ baʕdma
spilled-1S the-milk that Khaled felt-sick.3SM after

šerib -/=uh fi al-mṭṣam ʔms
drank.3SM-/=it in the-restaurant yesterday

I spilled the milk that Khaled became sick after he drank in the restaurant yesterday

2. šef-t al-jesr illi mḥammad safar qablma
saw.1S the-bridge that Muhammad travelled.3SM before
yhdem-un-/=uh fi al-ʔusbuṣ al-maḍi
destroy-3LPM-/=it in the-week the-last

I saw the bridge that Muhammad travelled before they destroyed in the last week

3. əštrai-t al-lawḥah illi aṭ-ṭullab inbahar-uu
bought-1S the-painting that the-students astonished-3PLM
lamman šaf-au-/=ha mṣallaqah ʕla jedar al-madrasah
when saw-3PLM-/=it hangng on wall the-school

I bought the painting that the students were astonished when they saw hanging on the school wall.

4. ʔkal-t al-ḥala illi mḥmmad inbaṣaṭ lamman
ate-1S the-dessert that Muhammad felt-happy.3M when
jeb-t-/=uh men al-maṭṣam al-jdeed
brought-1S-/=it from the-restaurant the-new.

I ate the dessert that Muhammad became happy when I brought from the new restaurant.

5. əštrai-t al-ketab illi ʔakhšyyat muna teḥassan-at
bought-1S the-book that personality Muna improved-3SF

lamman qra?-t/=uh khelal thlath šhuur
when read-3SF-/=it in three months

I bought the book that your personality improved when you read in three months

6. reḥ-t al-estrahah illi saleḥ raḥ yeʕzem aṣḥabuh
went-1S the-lounge that Saleh will invite.3SM frinds-his
law estʔjar-/=ha men ṣaḥeb-ha al-metkabber
if rent.3SM-/=it from owner-it the-arrogant

I went to the lounge that Saleh would invite his friends if he could rent from its arrogant owner

Crossed RC-Island:

A. Bare WH-questions:

1. ʔayš Mḥmmad ʕazam al-fannan illi rasam-/=uh
what Muhammad invited.3SM the-artist who drew.3SM -/=it
ʕla jedar al-maḥaf?
on wall the-museum?

What did Muhammad invite the artist who drew on the museum wall?

2. ʔayš saleh khaṣam al-sabbak illi aštra-/=h
what Saleh know.3SM the-man that bought.3SM-/=it
men mḥal as-sebaka?
from shop the-plumbing?

What does Saleh know the man who bought from the plumbing shop?

3. ʔayš khaled qabal al-mukhrej illi akhraj-/=uh fii
what Khaled met.3SM the-director that direct.3SM-/=it in

dar-al-ubera al-mşryyah?

house-the-opera the-Egyptian?

What did Khaled meet the director who directed in the Egyptian Opera?

4. ʔayš barak-ti li-lbent illi ektaşaf-at-/uh
what congratulate.2SF the-girl that invented-3SF-/it
min madrasat Raghadan?
from school Raghadan?

What did you congratulate the girl who invented from Raghadan school?

5. ʔayš ali şakar al-muhandis illi şammam-/uh
what Ali thanked.3SM the-engineer who designed.3SM-/it
fii muddah qaşeerah?
in period small

What did Ali thank the engineer who designed in small period?

6. ʔayš aş-şurṭah mesik-u al-ḥarami illi
what the-police arrested-3PLM the-thief who
eşraq-/uh min baitu-kum al-jdeed?
stole.3SM-/it from house-your the-new?

What did the police arrest the thief who stole from your new house?

B. Relative Clauses

1. ʔkal-t al-ḥala illi muna teʔref al-bent illi
ate-1S the-dessert that Muna knew.3SF the-girl that
saww-at-/uh fii al-ḥaflah ʔms
made-3SF-/it in the-party yesterday

I eat the dessert that Mona knows the girl who made in the party yesterday

2. mḥammad yudrus fii al-madrasah illi saleḥ yeṣref
Mhammad study.3SM in the-school that Saleh know.3SM
ar-jjal illi yamlek-/=ha fii ḥayy aš-sšafa
the-man who owned.3SM-/=it in district the-Shefa
Muhammad is studying at the school that Saleh knows the man who own in Shafa district.

3. habbai-t al-fellah illi qabl-na ar-rejjal illi šammam-/=ha
liked-1S the-vella that met-we the-man who designed.3SM-/=it
Ḥla kurnaiš jeddah
on beach Jeddah
I liked the vella that we met the man who designed on Jeddah beach

4. šreb-t al-Ḥaṣeer illi ahmad tsamm baḥd-ma
drank-1S the-juice that Ahmad poisoned.3M after
šerb-/=uh fii al-qhwah ḥms
drank.3SM-/=it in the-café yesterday
I drank the juice that Ahmad was poisoned after he drank in the café yesterday

5. dawwa-rt Ḥla al-ketab illi neṣref at-ṭaleb illi
lookerd-1S for the-book that know.1PL the-student who
əstaḤar-/=uh men al-ḥusbuḤ al-maḍi
borrowed.3SM-/=it from the-week the-last
I looked for the book that we know the student who borowed in the last week

6. ḥub al-ḥuḡniah illi muna teṣref al-muṭreb
liked-1S the-song that Muna know.3SF the-singer

illi laḥan-/=ha li-mhrajant al-ḡuḡniah al-ḡrabyyah
who wrote.3SM-/=it for-festival the-song the-Arabic

I like the song that Muna knows the singer who wrote for the Arabic Song festival

Discarded Experimental Sentences in Study 2

Bare Wh-Questions:

A. Short + Non-Island + Gaps (Adjunct Chain)

1. bi-ḡayš arsal-at muna khabar ḡenn-ha raḡ taftaḡ
By-what sent-3SF Mona news that-her w ill open
ṣaloon tajmeel?

salon beauty?

by what did Muna spread the news that she will open a beauty salon?

2. ḡala ḡayš katab-ti ḡnn al-muḡaḡarah tḡjalat
On what wrote-2SF that the-lecture postponed
li-bukrah?

to-tomorrow?

on what did you write that the lecture is postponed to tomorrow?

3. men ḡayš astantaj-at amal ḡnn jaru-hum kan ḡameel
From what concluded-3SF Amal that neighbour-them was agent
serrii

secret?

From what did Amal conclude that their neighbour was a secret agent?

4. ḡal ḡayš ḡḡlan-ti ḡnnu ma-raḡ yukuun fiiah
On what announced-2SF that no-will be there

muʔtamar laġawyyat hathi as-sanah?

conference linguistics this the-year?

On what did you announce that there will be no linguistics conference this year?

5. bi-ʔayš arsal-at maha šakwa ʔnn-ha wajah-at

By-what sent-3SF Maha complaint that-her faced-3SF

muškelah fi tašġeel al-barnamaj al-jadeed?

problem in operating the-program the-new?

By what did Maha send the complaint that she faced a problem in operating the new program?

6. ʕala ʔayš našar-ti ʔʕlan ʔnn raḥ ykuun

On what submitted-2SF announcement that will be

fiih waḥayf mutahah bi-l-qesm?

there jobs available in-the-department?

On what did you make an announcement that there will be available jobs in the department?

B. Short + Non-Island + RPs (Adjunct Chain)

1. ʔayš arsal-at muna b=uh khabar ʔenn-ha raḥ taftaḥ

what sent-3SF Mona by=it news that-her will open

šalon tajmeel?

salon beauty?

by what did Muna spread the news that she will open a beauty salon?

2. ʔayš katab-ti ʕalay=h ʔnn al-muḥaḍarah

What wrote-2SF on=it that the-lecture

tʔjalat li-bukrah?

postponed to-tomorrow?

on what did you write that the lecture is postponed to tomorrow?

3. ʔayš astantaj-at amal menn=uh ʔnn jaru-hum
What concluded-3SF Amal from=it that neighbour-them
kan ʕameel serrii
was agent secret?

From what did Amal conclude that their neighbour was a secret agent?

4. ʔayš ʔʕlan-ti ʕalay=h ʔnnu ma raḥ yukuun
What announced-2SF on=it that no will be
fiih muʔtamar laḡawyyat hathi as-sanah?

there conference linguistics this the-year?

On what did you announce that there will be no linguistics conference this year?

5. ʔayš arsal-at maha bu=h šakwa ʔnn-ha wajah-at
What sent-3SF Maha by=it complaint that-her faced-3SF
muškelah fi tašgeel al-barnamaj al-jadeed?
problem in operating the-program the-new?

By what did Maha send the complaint that she faced a problem in operating the new program?

6. ʔayš našart-l ʕalay=h ʔʕlan ʔnn raḥ ykuun
What submitted-2SF on=it announcement that will be
fiih waḏayf mutahah bi-l-qesm?

there jobs available in-the-department?

On what did you make an announcement that there will be available jobs in the department?

C. Short + Island + Gaps (Adjunct Chain)

1. فإنا ؟ayš ?ah mhammad lamman ruḥ-tum

On what fell-down.3SM Mhammad when went-2PL

al-ḡabah

the-forest

On what did Muhammad fall down when you were in the forest?

2. فان ؟ayš sʔal-tum al-mudeerah qablma tudkhl-un

About what asked-2PL the-manager before attend-2PL

?jtemaʕ al-ʔedarah?

meeting the-administration?

About what did you ask with the manager before you attended the administration meeting?

3. فان ؟ayš sʔal-ti muna lamman qabal-na-ha

About what asked-2SF Mona when met-1PL-her

fii al-kafee shop?

in the-coffee shop?

About what did you ask Mona when we met her in the coffee shop?

4. taḥt ?ayš dakhhal ath-thuʕban gablma yuqtal-uk

Under what went.3SM the-snake before killed.3SM=it

ṣaleḥ?

Saleh?

Under what did the snake go before Saleh killed it?

5. فإا ؟ays khall-at sara al-mfateeh gablma truḥ
On what left-3SF Sara the-keys before went.3SF
al-jameʕah?
the-university?
On what did Sara leave the keys before she went to the university?

6. fii ؟ays ṭalaʕ-tum lamman kun-tii maʕa ṣaḥbat-ek
In what watch-2PL when were-2S with friends-your
fii as-senima?
in cinema?
what did you watch when you were with your friend in cinema?

D. Short + Island + RPs (Adjunct Chain)

1. ؟ays ṭaḥ mḥammad ʕalay=h lamman ruḥ-t-um
What fell-down.3SM Mhammad on=it when went-2PL
al-ġabah?
the-forest?
On what did Muhammad fall down when you were in the forest?

2. ؟ays sʔal-tum al-mudeerah ʕann=uh qablma tudkhl-un
What asked-2PL the-managera bout=it before attend-2PL
ʔjtemaʕ al-ʔedarah?
meeting the-administration?
About what did you ask with the manager before you attended the
administration meeting?

3. ؟ays sʔal-ti muna ʕann=uh lamman qabal-na-ha

What asked-2SF Mona about-it when met-1PL-her
fi al-kafee shop?
in the-coffee shop?

About what did you ask Mona when we met her in the coffee shop?

4. ?ayš dakhal ath-thuṣban taḥt=uh gabma yuqṭal-uh
What went.3SM the-snake under-it before killed.3SM=it
ṣaleḥ?

Saleh?

Under what did the snake go before Saleh killed it?

5. ?ayš khall-at sara al-mfateeh ṣlaay=h gablma
What left-3SF Sara the-keys on-it before
truḥ al-jameʿah?
went.3SF the-university?

On what did Sara leave the keys before she went to the university?

6. ?ayš ṭalaṣ-tum fii=h lamman kun-tii maʿa ṣaḥbat-ek
What watch-2PL in=it when were-2SF with friends-your
fii as-senima?

in cinema?

what did you watch when you were with your friend in cinema?

Relative Clauses:

A. Short + Non-Island (Gap/RP)

1. ḥabbai-t al-spagtti illi amal saww-at-/=ha
Liked-1S the-Spaghetti that Amal made-3SF-/=it
li-ṣaḥbat-ha men al-madrasah al-thanawyyah

for-friends-her from the-school the-high

I liked the spaghetti that Amal made for her friends from the high school

2. rekeb-t ad-darrajah illi ṣaleḥ jab-/=ha
Rode-1S the-bicycle that Saleh brought.3SM-/=it
li-mḥammad fii al-ʿeed.
for-mohammad in the-Eid.

I rode the bicycle that Saleh brought for Muhammad in Eid.

3. ḥaḍar-t al-musabaqah illi amal naẓm-at-/=ha
Attended-1S the-competition that Amal organized-2SF-/=it
bain mdares al-baḥa al-ʿebtedaʿyyah
between schools al-baha the-primary.

I attended the competition that Amal organised between Al-baha's primary schools

4. sakan-t fii aš-šuggah illi ʔbo-yah ʔthath-/=ha
Stayed-1S in the-flat that father-my furnished.3SM-/=it
qabl zawaj mḥammad
before marriage Muhammad.

I stayed in the flat that my father furnished before Muhammad's marriage

5. ʔstakhdam-t al-komputer illi mona parmaj-at-/=uh
Used-1S the-computer that Mona programmed-3SF-/=it
qabl warṣat al-ʿamal
before the-workshop.

I used the computer that Maha programmed before the workshop.

6. mḥammad kasar as-saḥah illi amal jab-at-/=ha
Mhammad broke.3SM the-watch that Amal brought-3SF-/=it
li-ʔboy-ah fii ḥaflat at-taqaḥud
for-dad-my in party the-retire
Mohammad broke the watch that Amal brought for my father in the retirement party.

Sentences included in Study 3

Non-Crossed Island:

A. Non-D-linked WH-questions:

1. Context: Amal travelled to most of the cities in Saudi Arabia the last summer

ʔayš amal zar-t-/=h lamman kan-at fi ar-ryaḍ?

What Amal visit-3SF-/it when went-3SF in the-Riyadh?

What did you visit (it) when you went to Riyadh?

2. Context: Muhammad has studied hard for week

ʔayš raḥ taṣṭ-in-/=h mammad law nejeḥ fi al-ekhtebar?

what will give-2SF-/=it Muhammad if passes.3SM in the-exam?

What would you give Muhammad if he passed the exam?

3. Context: Muhammad went shopping in the morning with his friends.

ʔayš kan mḥammad yaḥtaj-/=uh qablma ysafer jeddah?

what was Muhammad need.3SM-/=it before travel.3SM Jeddah?

What did Muhammad need before he travelled to Jeddah?

4. Context: The pupils in my literature class are all very motivated this year.

ʔayš raḥ amal tulq-i-/=h law šarak-at

what will Amal recite.3SF-/=it if participated-3SF

fii al-musabaq al-ʔdabyyah?

in the-competition the-literature?

What would Amal recite if she participated in the literature competition?

5. Context: I'm sorry my children are so unruly.

ʔayš khaled kasar-/=uh lamman tḍarab maḥ

what Khaled broke.3SM-/=it when fought.3SM with

aṣḥa-buh?

friends-his?

What did Khaled break when he fought with his friends?

6. Context: My sisters will attend Amal's wedding party.

ʔayš raḥ afnan talbas-/uh law raḥ-at

what will Afnan wear.3SF-/=it if went-3SF

az-zawaj?

the-wedding-party?

What will Afnan wear if she attend the wedding party?

B. D-linked WH-questions:

Context:

1. Context: Most of my sisters love Asian food

ʔayy maṭṭam šini ʔkhwat-ek za-raw-/h

which restaurant Chinese sisters-your visted-3PL-/=it

lamman kanaw fii jeddah?

when were in Jeddah?

Which Chinese restaurant do your sisters visit when they go to Jeddah?

2. Context: I can't decide on the accommodation that fit my family when they come to Riyadh

ʔayy bayt muna astʔjar-at-/uh lamman ʔhla-ha

which house Muna rented-3SF -/=it when family-her

zar-aw-ha?

visted-3PL-her?

Which house did Mona rent for her family when they visited her?

3. Context: Muhammad goes to the hospital constantly to get physical therapy

ʔayy musabaqah mḥammad kheser-/=ha bisabab al-eṣabah?

which competition Muhammad lose.3SM -/=it because the-injury?

Which competition did Muhammad lose after being injured?

4. Context: The school asked the participants to be well-dressed

ʔayy badla khalid lebes-/=ha lamman šarak

which suit Khaled wore.3SM -/=it when participated.3SM

fii al-musabqah?

in the-competition?

Which suit did Khaled wear when he participated in the competition?

5. Context: This writer used to write about sensitive topics

ʔayy maqal al-kateb našar-/=uh qablma

which article the-author published.3SM -/=it before

yastaqeel min aṣ-ṣaḥifah?

retired.3SM from the-newspaper?

Which article did the author publish before he retired from the newspaper?

6. Context: I think that there will be an increase in the number of car parking after ending the project

ʔayy mbani al-baladyah raḥ tšil-/ha lamman

which building the-council will remove.3SF -/=it when

ywasseʔ-uun manṭeqat mwaqef as-sayyarat?

expand-3PLM area parking the-cars?

Which buildings will the council remove when they expand the parking area?

Crossed Non-Island:

A. Non-D-linked WH-questions:

1. Context: I don't know when I can visit Muhammad to congratulate him for his new business

ʔayš tetwaqʕ-een ʔnn mħmmad raħ yebeeʕ-/=h fii
what think-2SF that Muhammad will sell.3SM-/=it in
mħall-uh al-jdeed?
shop-his the-new?

What do you think that Muhammad will sell in his new shop?

2. Context: The art class is very engaging for the students

ʔayš yabġa khaled yarsum-/=uh ʕla jedar
what want.3SM Khaled draw.3SM-/it on wall
al-fašl?
the-classroom

What does Khaled want to draw on the classroom wall?

3. Context: The faculty staff complains about many problems in the department

ʔayš tetwaqaʕ-een inn al-mudeerah raħ tenaqiš-/=uh
what think-.2SF that the-adminstrator will discuss.3SF-/=it
fii ejtemaʕ al-edarah?
in meeting the-administration?

What do you think that the administrator will discuss in the administration meeting?

4. Context: People of Al-baha said that the government cares about improving their city.

ʔayš qal aħmad inn al-baladyyah raħ tebni-/=h
what said.3SM Ahmad that the-council will build.3SM-/=it

fii wast al-madiinah?

in centre the-city?

What did Ahmad say that the council will build in city centre?

5. Context: I am impressed with the decorations of Khaled's new house

?ayš tetwaqaʕ=een inn ar-rassam bi-yarsum-/uh

what think-2SF that the-artist will-draw.3SM-/=it

ʕla al-jedar al-khareji?

on the-wall the-outer?

What do you think that the artist will draw on the outer wall?

6. Context: The girls who attended the wedding party noticed that Mona was jealous that her sister is getting married

?ayš tetwaqaʕ-un inn maha bi-tquul-/uh baʕd

what think-3PL that Maha will-say.3SF-/=it after

al-ḥflah az-zawaj?

party the-wedding?

What did Maha say that she will bring after the wedding party?

B. D-linked WH-questions:

1. Context: Muhammad's car is still in the car workshop

?ayy sayyarah khaled qal inn mḥammad

which car Khaled said.3SM that Muhammad

ṣadam-/=ha fii ṭreeq-uh li-l-ḡabah?

hit.3SM -/=it in way-his to-the-forest?

Which car did Khaled say that Muhammad hit on his way to the forest?

2. Context: Muhammad went to the supermarket to buy some stuff for the party

ʔayy ʔala mʔammad qal inn ʔkht-uh raʔ
which dessert Muhammad said.3SM that sister-his will
ʔjahez-/uh li-ʔʂʔab-uh fi al-layl?
prepare.3SF-/=it for-friends-his in the-night?

Which dessert did Muhammad say that his sister prepares for his friends at night?

3. Context: My brother is still working in the engineering office in Makkah

ʔayy vella qul-ti l-e inn akhu-k
which palace told-2SF for-me that brother-your
ʂamma-/=ha li-mudeer ʂarekat al-esment?
designed.3SM-/=it for-manager firm the-cement?

Which palace did you tell me that your brother designed for the cement firm's manager?

4. Context: I want to watch an action movie this weekend

ʔayy film tagreed gal-at inn-ha tabaʕ-at-/=uh
which movie Tagreed said-3SF that-she watched-3SF-/=it
mʕa ʂaʔbatha ʔms?
with friends-her yesterday?

Which movie did Tagreed say that she watched with her friends yesterday?

5. Context: Mona passed all the exams of English language with excellent marks

ʔayy kitab muna qal-at enn-ha ʔakar-at-/=uh qabl
which book Muna said-3SF that-she studied-3SF-/=it before
ekhtebar al-IELTS?
exam the-IELTS?

Which book did Mona say that she studied before the IELTS exam?

6. Context: Fatma sent me an invitation for her wedding party

ʔayy fustan tabġ-een talbas-een-/uh eġa reḥ-ti
which dress want-2SF wear-2SF-/=it if attend-2SF
az-zawaj?
the-wedding-party?

Which dress do you want to wear if you attended the wedding party?

Crossed Adjunct-Island:

- A. Non-D-linked WH-questions:

1. Context: Mona's schedule was busy yesterday. Even, she had no time for reading to get ready for the lecture.

ʔayš muna ḥaḍar-at al-muḥaḍarah qabl-ma taqr?-/=h
what Muna attended-3SF the-lecture before read.3SF -/=it
fii maktabat al-jameʕah?
in library the-university?

What did Mona attended the lecture before she read in the university library?

2. Context: My result wasn't good because I understood too late what the questions were about.

ʔayš kammal-ti ḥal al-ʔsʔelah bidunma
what finished-2SF answering the-questions without
tlaḥẓeen-/uh fii nehayat al-ekhtebar?
notice-2SF-/=it at end the-exam?

What did you finish answering the questions before you noticed at the end of the exam?

3. Context: Praise be to Allah that most house' rooms were closed when the thief was inside the house.

ʔayš aš-šurṭah mesik-uu al-ḥarami lamman
what the-police arrested-3PLM the-thief when
əstraq-/uh min baitu-kum al-jdeed?
stole.3SM-/it from house-your the-new?

What did the police arrest the thief when he stole from your new house?

4. Context: The employees who were careless became hardworking to avoid being fired in the same way as their friend.

ʔayš al-mudeer faṣal al-muḏaf lamman
what the-administrator sacked.3SM the-employee when
ahmal-/uh fii akher jahrain?
dismissed.3SM-/it in last month.DU?

What did the administrator sack the employer when he dismissed in the last two months?

5. Context: Muhammad told the engineer that the company is planning to build a restaurant with a luxurious design

ʔayš ahmad ekhtar al-muhandes baʕdma ṣammem-/uh
what Ahmad chose.3SM the-engineer after designed.3SM-/it
fii mul al-ʕrab?
in mall the-Arab?

What did Ahmad choose the engineer after he designed in Arab Mall?

6. Context: The children who attended yesterday's party were naughty

ʔayš maha zeʕel-at lamman amal kasar-at-/uh
what Maha felt-angry-3F when Amal broke.3F-/it

fii ḥaflat ?ms?

in party yesterday?

What did Maha become angry when Amal broke in yesterday's party?

B. D-linked WH-questions:

1. Context: The social worker came to visit Ahmad in prison yesterday

?ayy juruṭ raḥ yefrej-uun ṣan aḥmad eḍa

which conditions will release-3PLM about Ahmad if

ejtaz-/=ha fii settat ?šhur?

passed.3SM-/=it in six months?

Which conditions will they release Ahmad if he met after six months?

2. Context: I have to buy a new mobile phone this month

?ayy šaḥen khereb jawwal-ek al-qadeem

which charger damaged mobile-your the-old

lamman estakhdam-ti-/=h fii al-cafe ?ms?

when used-2SF-/=it in the-café yesterday?

Which charger did your old mobile damage when you used in the café shop yesterday?

3. Context: My sisters want to go to the tailor this week

?ayy fustan kholoud raḥ tuḥḍur az-zawaj law

which dress Kholoud will attend.3SF the-wedding if

mada-ha tṣaddel-/=uh qabl yawm al-?ḥad?

could-her alter.3SF-/=it before day the-Sunday

Which dress will Kholoud attend the wedding party if she could alter before Saturday?

4. Context: The art show was special this year

ʔayy lawḥah aṭ-ṭullab enbaha-ruu lamman
which painting the-students astonished-3PLM when
ʃafa-u-/=ha mʃallaqah ʃla jedar al-madrasa?
saw-3PLM-/=it hanging on wall the-school?

Which painting were the students astonished when they saw hanging on the school's wall?

5. Context: I like to take lots of books with me when I travel.

ʔayy ketab ḥassai-ti inn šakḥṣyyat-ik teṭawwar-at
which book feel-2SF that personality-your improved.3SF
lamman qarʔ-ti-/=h fii al-ʔjazah al-maḍiah?
when read-2SF-/=it in the-holiday the-last?

Which book do you feel that your personality improved when you read the last holiday?

6. Context: The administrator honored Maha in the school

ʔayy qaṣeedah amal faz-at lamman ʔlq-at-/=ha
which poem Amal won-3SF when recited-3SF -/=it
fi musabaqat al-ʔdab al-ʃrabi?

In competition the-literature the-Arabic?

Which poem did Amal win when she recited in the Arabic literature competition?

Crossed RC-Island:

- A. Non-D-linked WH-questions:

1. Context: Muhammad met the artists who are working with my father's company in Jeddah

ʔayš mḥmmad ʃazam al-fannan illi rasam-/=uh

what Muhammad invited.3SM the-artist who drew.3SM -/=it

ʕla jedar al-maḥaf?

on wall the-museum?

What did Muhammad invited the artist who drew on the museum wall?

2. Context: Saleh did not know from where he can buy some building material, so he depended on plumbers to bring them for him

ʔayš saleh yeʕref al-sabbak illi jab-/=h

what Saleh know.3SM the-plumber that bought.3SM-/=it

men mḥal as-sebaka?

from shop the-plumbing?

What does Saleh know the man who bought from the plumbing shop?

3. Context: Khaled has lots of contacts in the music world.

ʔayš khaled qabal al-mukhreḥ illi akhraj-/=uh

what Khaled met.3SM the-director that direct.3SM-/=it

fii dar-al-ubera al-mṣryyah?

In house-the-opera the-Egyptian?

What did Khaled meet the director who directed in the Egyptian Opera?

4. Context: We met the students who were qualified for the talented students competition from Al-baha schools.

ʔayš barak-ti l-bent illi ektašafat-/uh min

what congratulate.2SF the-girl that invented-3SF-/=it from

madrasat raghadan?

school Raghadan?

What did you congratulate the girl who invented from Raghadan school?

5. Context: This company has a very strange staffing policy, that doesn't take efficiency into account.

ʔayš al-mudeer šakar al-muhandis illi
what the-manager thanked.3SM the-engineer who
šammam-/uh fii muddah qašeerah?
designed.3SM-/it in period small?

What did the manager thank the engineer who designed in small period?

6. Context: Praise be to Allah that the police car was in the neighborhood.

ʔayš aš-šurṭah mesik-u al-ḥarami illi
what the-police arrested-3PLM the-thief who
əstraq-/uh min baitu-kum al-jdeed?
stole.3SM-/it from house-your the-new?

What did the police arrest the thief who stole from your new house?

B. D-linked WH-questions:

1. Context: Policemen are everywhere to reduce disorder

ʔayy sayyarah al-ḥarami qatal al-walad illi
which car the-thief killed-3SM the-boy who
kan ysuq-/ha fii wasaṭ al-madeenah?
was.3SM driving.3SM-/it in centre the-city?

Which car did the thief kill the man who was driving in the city centre?

2. Context: Saleh has a lot of contacts in the construction world

ʔayy mabna saleh yeṣref al-muqauel illi
which building Saleh know.3SM the-contractor who
rammam-/ha fii ḥay aš-šafa?
repaired.3SM-/it in district the-Shafa

Which building does Saleh Know the contractor who repaired in Ash-Shafa district?

3. Context: I like to watch morning programs everyday

ʔayy qaʕeedah estɔafa-uu aʕ-ʕaʕer illi katab-/=ha
which poem host-3PLM the-poet who write.3SM-/=it
fii barnamj ʕabah al-khair
in program morning the-good.

Which poem did they host the poet who wrote in the (Good Morning Arabs) program?

4. Context: I liked the school's care of the talented students in drawing

ʔayy lawha al-mudeerah karram-at aṭ-ṭalebah
which painting the-administrator.F honored-3SF the-student
illi rasm-at-/=ha fii al-maʕrv al-fanni?
that draw-3SF-/=it in the-exhibition the-art?

Which painting did the administrator honor the girl that painted for the art exhibition?

5. Context: I always like to read the Carpathian articles in the newspapers

ʔayy jameʕah teʕrf-een al-kateb illi
which university know-2SF the-author that
intaqad-/=ha fii maqal-uh al-ʔusbuʕi?
criticized.3SM-/=it in article-his the-weekly?

Which university did you know the author who criticized in his weekly article?

6. Context: I am interested in reading the gossip of the celebrities in the music world.

ʔayy ʔuġnyah faz al-mulaḥen illi laḥan-/=ha
which song won.3SM the-poet who wrote.3SM-/=it
fii musabaqat al-ʔuġnyah al-Ṣarabyyah?
in competition the-song the-Arabic

Which song did the poet who wrote win the competition in the Arabic Song Contest?

Sentences included in Study 4

1. Salient context

Mona wanted to borrow the introductory phonetics book before attending the lecture. On the next day, she went to the college to borrow the book from the library. Unfortunately, she found that it is not available for borrowing.

Non-salient context

Mona was too busy to borrow the introductory phonetics book before attending the lecture. She did not have the time to go to the library in advance. As a result, she could not understand most of the new information in the lecture.

Experimental Sentence

ʔayy ketab / ʔayš muna ḥaḍar-at al-muḥaḍarah beduunma
which book/ what Muna attended.3SF the-lecture without
testeʕeer-/uh men maktabat al-jameʕah?
borrow.3SF -/=it from library the-university?

Which book/what did Mona attended the lecture without borrowing -/it from the university library?

2. Salient context

One of the children at yesterday's party broke Hind's Samsung mobile phone. Hind was angry because her phone was new and expensive. She had received it as a gift from her brother for her birthday.

Non-salient context

The children who attended yesterday's party were so naughty that one of them broke Hind's Samsung mobile phone. Hind was angry because the children had been so naughty. She did not expect her guests to be so disrespectful of her house.

Experimental Sentence

ʔayy jawwal / ʔayš hind zeʕel-at lamman al-waalad
which mobile-phone/what Hind felt-angry.3SF when the-boy
kasar-/uh fii ḥaflat ʔms?
broke.3SM -/=it in party yesterday?

Which mobile phone/what was Hind angry when the child broke -/it in yesterday's party?

3. Salient context

The engineer who designed the restaurants building in Hada Park won in the competition of the best architectural design. The building design was inspired by the mountain nature of the city. I intended to visit it next week with my family.

Non-salient context

The engineer who designed the restaurants building in Hada Park won in the competition of the best architectural design. The engineer has just earned his master degree in architectural engineering from Jeddah University. He is considered one of the best architectural engineers in the area.

Experimental Sentence

ʔayy mabna / ʔayš al-muhandes faz fi al-musabaqah
which building/ what the-engineer won.3SM in the-competition
baʔdma šammam-/uh fii muntazah al-hada?
after designed.3SM -/it in park the-Hada?

Which building/what did the engineer win in the competition after he designed -/it in Hada Park?

4. Salient context

Amal intended to wear the black dress in her friend's wedding party. She got the dress from the tailor one week before the wedding party. But since she gained weight during that time, it didn't fit her perfectly.

Non-salient context

Amal didn't wear the black dress in her friend's wedding party because she gained weight. She had lots of exams along the last two months, so she stopped going to the gym. Approximately, she gained about 4 kg.

Experimental Sentence

ʔayy fustan/ ʔayš amal meten-at qablma talbas-/uh
which dress/ what Amal gain.wieght-3SF before wear.3SF -/it
fi zawaj šaḥbat-ha?

in wedding-party friend-her?

Which dress/what did Amal gain weight before she wore _/it in her friend's wedding party?

5. Salient context

Amal was poisoned after eating the grilled prawns dish in the lounge buffet. The dish was spicy, and the prawns were perfectly cooked. But unfortunately, it seems to be poisoned.

Non-salient context

Amal was poisoned after eating the grilled prawns dish in the lounge buffet. More than one person was poisoned yesterday. The council closed the restaurant that was responsible for the buffet, and fined its owner.

Experimental Sentence

ʔayy	ṭabaq/	ʔayš	amal	tsammam-at	baʕd-ma	ʔkal-at-/uh
which	dish/	what	Amal	poisoned-3SF	after	ate-3SF -/=it
fii	bufayh		al-estrahah?			
in	buffet		the-lounge			

Which dish/what was Amal poisoned after she ate _/it in the lounge buffet?

6. Salient context

Mona got emotional in yesterday's graduation party after listening to a pathos poem recited by a student for her passed away father. The poem was describing the suffering of the girl after the death of her father. The girl was so emotional while reciting it.

Non-salient context

Mona got emotional after listening to a pathos poem yesterday. But generally, she enjoyed most of the parts of the celebration. There were competitions, songs and mimes, and the attendees were actively engaged with each part of the celebration.

Experimental Sentence

ʔayy	qaṣeedah/	ʔayš	maha	tʔθr-at	lamman
which	poem/	what	Maha	felt-emotional.3SF	when
semʕ-at-/ha		fii ḥafl		at-takharruj	ʔms?

listened-3SF -/=it in party the-graduation yesterday?

Which poem/what was Maha emotional when she listened to -/it in yesterday's graduation party?

7. Salient context

Ahmad read the new English textbook with the private teacher before the exam. The book contains exercises that illustrate the strategies for answering the exam questions. Thank God, Ahmad benefited from it.

Non-salient context

Ahmad passed his exam after studying the new English textbook with the private teacher. Now, he wants to register in workshops to improve his computer skills. In particular, he wants to improve those skills that are related to data analysis.

Experimental Sentence

ʔayy ketab/ ʔayš ahmad tjawaz al-ekhtebar baʕd-ma
which book/ what Ahmad passed.3SM the-exam after
qraʔ-/=uh maʕ al-mudarres al-khuṣuṣi?
read.3SM-/=it with the-teacher the-private?

Which book/what did Ahmad pass IELTS after he studied _/it with the private teacher?

8. Salient context

The lotus sweet of this sweet shop is outstanding and delicious. But unfortunately, when I went to the shop to buy some of the sweet for Khaled's birthday party, I found the shop closed. If I could bring some of it, Khaled would be so happy.

Non-salient context

When I went to the sweet shop to buy the lotus sweet for Khaled's birthday party, I found the shop closed. Although we left home early, there was a traffic congestion. We waited about an hour for the traffic congestion to be reduced

Experimental Sentence

ʔayy ḥala/ ʔayš gaffal al-mḥal qablma teṣṭr-eeen-/=uh
which sweet/ what closed.3SM the-shop before buy-2SF -/=it
le-ḥaflat milad khaled?

for-party birthday Khaled?

Which sweet/what was the shop closed before you buy -/it for Khaled's birthday party?

9. Salient context

Sara used mebo ointment on the burn scars in her left hand. Thanks God, the scars disappeared completely after she used this ointment. She advises anyone who have burn scars to use it

Non-salient context

The burn scars in Sara's left hand disappeared completely after she used memo ointment. Hot coffee spilled over her hand while she was serving it to guests. But, thanks God, most of the burns were superficial.

Experimental Sentence

ʔayy	kream/	ʔayš ʔaθar al-ħuruuq ekhtaf-at	lamman
which	ointment/	what scars the-burns disappeared-3SF	when
sara	estakhdam-at-/uh	ʕla yedda-ha al-yesar?	
Sara	used-3SF -/=it	on hand-her the-left?	

Which ointment/what did the burn scars disappear when Sara used -/it on her left hand?

10. Salient context

Dr Ahmed carried out a study on a new cancer treatment. It took him about two years to complete the study. After publishing it in a well-known international journal, the doctor was promoted in his work

Non-salient context

Dr Ahmad was promoted after publishing a research on a new cancer treatment in a well-known journal. The university used to encourage staff members to consider publishing in well-known international journals. Indeed, the university ranking improved the last year.

Experimental Sentence

ʔayy	baħθ/	ʔayš traqqa	al-ʔustað	lamman
which	research/	what promoted.3SM	the-doctor	when

našar-/uh fi al-majallah al-ʕelmyyah?

published.3SM -/it in the-jurnal the-international?

Which research/what was the doctor promoted after he published _/it in an international journal?

11. Salient context

Mona used Ahmad's computer and deleted his most recent folder by mistake. Ahmad was so nervous when he knew, as his latest work was saved in that folder. Fortunately, he could recover it back later

Non-salient context

Ahmad was nervous when he knew that Mona deleted his most recent folder from his computer. He didn't expect that Mona will use his computer without his permission. He wants to teach her a lesson that she can't interfere in others' affairs.

Experimental Sentence

ʔayy malaf/ ʔayš ahmad zeʕel lamman muna

which folder/ what Ahmad felt-angry.2SM when Muna

ḥaḏaf-at-/uh men jehaz al-computer?

deleted-3SF-/it from machine the-computer?

Which folder/what was Ahmad nervous when Mona deleted -/it from the computer?

12. Salient context

Mona was excited to attend a presentation a presentation on technology in education in the conference. Unfortunately, the doctor who was meant to deliver the presentation apologised for not attending the conference. The organisers were obliged to remove it from the conference program.

Non-salient context

Mona was upset to know that the organisers removed the technology in education presentation from the conference program. But generally, the conference was useful and the program was diverse. There were participations from well-known authors in the field.

Experimental Sentence

ʔayy ʕarḏ / ʔayš muna zeʕl-et lamman

which presentation/ what Muna felt-angry-3SF when
al-munaẓmeen alġa-uu-/=h men parnamaj al-mu?tamar?
the-organizers removed-3PL -/=it from program the-conference?

Which presentation/what was Mona upset when the organisers removed _/it from the conference program?

b- How do you learn your second language?

-At home _____

-In school _____

-Living in a country where the second language is spoken _____

c- Please specify the age at which you started to learn your second language in the following situations (write age next to any situation that applies).

At home: _____

In school: _____

After arriving in the second language speaking country: _____

d- In general, how would you rate your English language proficiency?

None _____

Poor _____

Fair _____

Good _____

Very good _____

Native-like _____

APPENDIX C Language Background Questionnaire (in Arabic)

١- الجنس: ا: ذكر ب: انثى

٢ العمر (بالسنوات) _____

٣ ماهي اللغة الاصلية (اللغة الام) التي تتحدث/ين بها _____

٤ اذا كانت اللغة العربية هي لغتك الاصلية، فما هي اللهجة العربية التي تحدثت/ي بها في مرحلة الطفولة (مثال: اللهجة الحجازية/ اللهجة النجدية/ لهجة الباحة/ اللهجة المصرية/ اللهجة العراقية الخ).

اذا كنت/ي تتحدث/ين باكثر من لهجة، حدد/ي اللهجة التي استخدمتها اكثر. _____

٥ اي لهجة من اللهجات العربية تتحدث/ين بها في حياتك اليومية (مثال: اللهجة الحجازية/ اللهجة النجدية/

لهجة الباحة/ اللهجة المصرية/ اللهجة العراقية الخ)

اذا كنت/ي تتحدث/ين باكثر من لهجة، حدد/ي اللهجة الاكثر

استخداما. _____

٦- هل ولدت/ي في

الباحة؟

*اذا كانت اجابتك ب لا:

ا: ما هي المدينة التي ولدت/ي فيها؟ _____

ب: متى انتقلت/ي للعيش في

الباحة؟

٧- هل تتحدث اي لغات اخرى (مثال: الانجليزية/ الفرنسية/ الاسبانية الخ)

ا: نعم ب: لا

* اذا كانت اجابتك بنعم:

ا: ماهي هذه اللغات؟ _____

ب: كيف تعلمت/ي لغتك الثانية؟

١- في المدرسة

٢- في المنزل

٣- بعد العيش في دولة يتحدث سكانها هذه اللغة

ج: حدد/ي العمر الذي تعلمت/ي فيه اللغة الثانية (اكتبي العمر مقابل الاختيار المناسب)

١. في المدرسة

٢. في المنزل

٣. بعد العيش في دولة يتحدث سكانها هذه اللغة

د: كيف تقويم/ين مستواك في اللغة الانجليزية

١. سيء جدا

٢. سيء

٣. متوسط

٤. جيد

٥. جيد جدا

٦. كمتحدث اصلي للغة

٥: اذا كنت تتحدثين لغة اخرى غير العربية او الانجليزية، كيف تقيم/ين مستواك فيها

١. سيء جدا

٢. سيء

٣. متوسط

٤. جيد

٥. جيد جدا

٦. كمتحدث اصلي للغة

APPENDIX D Information Sheet (English Version)

Participant Information Sheet

Linguistics and Phonetics

Department University of Leeds

Study title: Understanding complex sentences Investigator: Asmaa Alghamdi

Introduction

I am a researcher in the Linguistics & Phonetics Department at the University of Leeds and I am interested in how native speakers of Baha Arabic read complex sentences and how this is related to their memory capacity. You have been invited to participate as you are a native speaker of Baha Arabic, which is the language under investigation in this study.

Before you decide to take part in this study it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and discuss it with others if you wish. Ask me if there is anything that is not clear or if you would like more information. Take time to decide whether or not you wish to take part.

The Experimental tasks

There will be two tasks in this experiment. In the first one, you will be presented with blocks which will be marked in a particular order and you will be asked to remember the exact order in which the blocks are marked. In the second task, you will be asked to read a number of sentences, which are presented word-by-word on a computer screen. In particular, by pressing a particular key, the first word in a sentence will appear, and by pressing it again, this word will disappear, and the next one will appear; i.e. each subsequent key-press reveals the next word in the sentence and hide the previous word. There will be a total of 96 sentences to read. After reading each sentence, you will be asked to either answer a yes/no question related to the previously read sentence, or to judge its acceptability.

Benefits and risks

There are no risks involved and by taking part in this study you will have the chance to be involved in real research. Participants will be awarded course credits for taking part in the research.

Withdrawal from study

You are free to withdraw from the experiment at any point without prejudice and without needing to give a reason. In addition, at any time during the testing, the investigators have the right to terminate the study for any reason.

However, participants cannot withdraw from the study once the results have been written up or published.

Confidentiality

Your identity will be kept anonymous. Participants will only be identifiable through a unique code, and any links between your identity and code will be stored in secure servers and/or hard disks secured by the University of Leeds network security procedures. You will be given the option to have your name kept on record if you wish to be contacted with regards to participation in future studies. Any data provided will only be accessible to the researcher and supervisors, although it is possible that your data may be used in future conference presentations or journal article publications; your identity will be disguised during this process. All information provided by you will be kept confidential.

If you have any further questions please ask me.

Contact

For further information contact Asmaa Alghamdi (ml14aama@leeds.ac.uk) or the supervisor of the PhD student Cecile De Cat (c.decat@leeds.ac.uk).

APPENDIX E Information Sheet (Arabic Version)

ورقة المعلومات للمشارك
قسم اللغويات و الصوتيات
جامعة ليدز

عنوان الدراسة: فهم الجمل الصعبة
الباحث: أسماء الغامدي

مقدمة:

انا باحثة في قسم اللغويات و الصوتيات بجامعة ليدز البريطانية و مهتمه بالبحث عن الكيفية التي يقرأ بها المتحدثون بلغة الباحة العربية الجمل الصعبة في لغتهم (الباحة العربية) و مدى ارتباط ذلك بالقدرة الاستيعابية للذاكرة. لقد تم اختيارك للمشاركة في هذا البحث لكونك متحدث بلغة الباحة العربية، وهي اللغة تحت الدراسة في هذا المشروع البحثي.

من الأهمية قبل مشاركتك في هذا البحث توضيح السبب من القيام بهذه الدراسة و ما ذا تتضمن. ارجو منك اعطاء الوقت الكافي لقراءة المعلومات الواردة ادناه و مناقشتها مع الاخرين اذا اردت ذلك. اسألني اذا ماكان هناك اي شيء غير واضح او اذا كنت تحتاج معلومات اكثر. خذ الوقت الكافي لتقرر اذا ما كنت ترغب بالمشاركة ام لا.

الواجبات في التجربة:

تتضمن هذه التجربة القيام بواجبين: في الواجب الاول سيتم عرض مجموعة من الجمل (مأخوذة من لغة الباحة العربية) واحدة تلو الاخرى على شاشة الكمبيوتر، و يطلب منك ان تقيم مدى امكانية حصول هذه الجمل في لغة الباحة العربية ، و سيكون مجموع الجمل المطلوب تقييمها ٤٣٢ جملة (مقسومة على استينانين). في الواجب الثاني، سيتم عرض مجموعة من الارقام على شاشة الكمبيوتر، وفي نهاية عرض كل مجموعة من الارقام سيطلب منك ادخال الارقام التي تم عرضها عليك في شاشة التذكر ولكن بترتيب معكوس.

الفوائد و المخاطر:

لن يكون هناك اي مخاطر في المشاركة في هذه التجربة ، و بمشاركتك في هذا البحث سيكون لك الفرصة بالمشاركة في بحث حقيقي.

الانسحاب من الدراسة:

لديك الحرية الكاملة في الانسحاب من التجربة في اي وقت بدون ضرر او الحاجة الى ابداء اي سبب. ايضا، يحق للباحث في اي وقت خلال التجربة ان يوقفها لاي سبب كان.

السرية:

هويتك سوف تبقى مجهولة، حيث انه سيتم تعريف كل مشترك برمز معين، و اي صلة تكون بين هوية المشترك و الرمز سوف تبقى محفوظة في سيرفر او اقرص صلبة محفوظة بواسطة شبكة الحماية في جامعة ليدز. كذلك، اذا كان لديك الرغبة ان يتم الاتصال بك مستقبلا للمشاركة في بحوث علمية، فانه لديك الخيار بان تبقى اسمك محفوظا في السجلات.

بالنسبة للبيانات التي قمت بادلائها، فانه سيكون لدى لباحث و المشرف القدرة على الدخول اليها.

كذلك فانه من الممكن استخدام بياناتك التي ادليت بها في عروض مؤتمرات قادمة او بحوث علمية منشورة،
و لكن ستبقى هويتك مجهولة في هذه الاحوال.
جميع المعلومات التي ستدلي بها ستبقى محفوظة بسرية تامة.

اذا كان لديك اي استفسارات اخرى، يمكنك التواصل على:
(ml14aama@leeds.ac.uk).

APPENDIX F Consent Form (English Version)

Thank you very much for your interest in our research. The purpose of this form is to make sure that you have been given a full and clear explanation of what is involved in the study, that you meet certain criteria, and that you are happy to take part.

	Add your initials next to the statement if you agree
I am over 18 years of age	
I confirm that I have read and understood the information sheet explaining the above research project and I have had the opportunity to ask questions about the project.	
I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason and without there being any negative consequences. In addition, should I not wish to answer any particular question or questions, I am free to decline.	
I give permission for members of the research team to have access to my anonymised responses. I understand that my name will not be linked with the research materials, and I will not be identified or identifiable in the report or reports that result from the research. I understand that my responses will be kept strictly confidential	
I agree that my data can be used for academic presentations and publications provided that my anonymity is maintained	
I agree to take part in the above research project.	

Name of participant	
Participant's signature	
Date	
Name of lead researcher	
Signature	
Date	

APPENDIX G Consent Form (Arabic Version)

عنوان الدراسة: فهم الجمل الصعبة
الباحث: أسماء الغامدي

في البداية اود ان اشكرك لمشاركتك في هذا البحث. الغاية من هذه الورقة هو التأكد من انه تم تقديم شرح كافي للمشارك عن ما تتضمنه الدراسة، و التأكد من ان المشترك يتمتع ببعض المقاييس المحددة، و انه موافق على المشاركة في الدراسة.

ضع الحرفين الاولى من اسمك اذا كنت موافق	
	عمري يتجاوز ال ١٨ عاما
	اقر بانني قرأت وفهمت ورقة المعلومات التي تشرح موضوع البحث و انني حصلت على الفرصة الكافية للسؤال عن هذا البحث
	افهم بان مشاركتي في البحث تطوعية و انه لي الحق في الانسحاب من المشاركة في اي وقت بدون اعطاء اي سبب و بدون ان يكون هناك اي عواقب سلبية. كذلك فانه لي الحرية الكاملة في الامتناع عن اعطاء اجابة لاي سؤال او اسئلة.
	اعطي الاذن لاي عضو في فريق البحث بالدخول الى اجاباتي والتي ستكون مجهولة المصدر. كذلك فانني افهم ان اسمي لن يكون مرتبطا بمحتويات البحث، وانني لن اكون معروفا او قابلا للتعريف في اي تقرير ينتج عن هذا البحث. افهم ان اجاباتي سوف تبقى محفوظة بسرية تامة.
	اوافق على انه بالامكان استخدام بياناتي في البحوث العلمية شريطة المحافظة على هويتي مجهولة.
	اوافق على المشاركة في البحث المذكور اعلاه.

	اسم المشترك
	التوقيع (يمكنك كتابة اسمك الاول فقط)
	التاريخ
	اسم الباحث
	التوقيع
	التاريخ