Accepting Preposition-Stranding under Sluicing Cross-linguistically; a Noisy-Channel Approach

Emilia Molimpakis

UCL

PhD Linguistics

I, Emilia Molimpakis, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

E. Molimpakis
(May, 2019)

Abstract

This thesis investigates the representation and processing of sluicing, a type of ellipsis where an interrogative CP is reduced to its initial wh-element (the remnant), e.g. *Mary danced with someone, but I can't remember (with) who.* It is debated whether remnants from within a PP (*with who*) must appear with this P or whether they can appear without it ('P-stranding'). Existing theoretical literature (Merchant, 2001; a.o.) argues that only languages allowing overt CPs to move *wh*-elements without their embedding P will allow P-stranding remnants (P-Stranding Generalisation/PSG). Anecdotally, many languages appear to defy this pattern, allowing P-stranding remnants despite disallowing P-stranding overtly. None of these examples, however, are supported by adequate experimental evidence, nor offer a cross-linguistically generalisable explanation. This thesis addresses both these issues. Novel large-scale acceptability data show that both Greek and German, previously proposed robust PSG-examples, do indeed defy it. This behaviour is explained by proposing ellipsis is a type of 'noisy channel' (Shannon, 1948; Gibson, Bergen & Piantadosi, 2013), through which the parser must estimate the probability of the intended (elided) message. The parser simultaneously considers the prior likelihood of the intended message (a remnant as part of a full PP) as well as the likelihood of this message being corrupted through 'noise' (a deleted P). P-stranding is thus considered a form of deletion, given deletion has been shown to be a likely corruption in noisy channels. A series of reading time studies aimed at supporting this noisy channel model in online processing found results overall consistent with this approach, but also discovered previous work on the processing of sluicing was inaccurate in concluding its active prediction by the parser. Collectively, the work argues for a theory of sluicing involving syntactic structure at the *e*-site together with sluicing being treated as a noisy channel by the parser.

Impact Statement

This thesis investigates the topic of ellipsis, a unique linguistic phenomenon that is prolific across all languages. Under ellipsis (1a), meaning is successfully transmitted in the absence of overt linguistic input (i.e. (1b) is understood).

(1) a. Someone was at the door yesterday, but I don't know who.
    b. Someone was at the door yesterday, but I don't know who was at the door yesterday.

All language learners succeed in comprehending ellipsis, despite the absence of overt data to guide them. This indicates that the learner must somehow recover meaning from the sentential context. This process has been argued to reflect innate cognitive biases driving not only language learning, but successful information conveyance, more generally. Be it during learning or in the adult processor, the key question is what contextual information is retrieved and how to fill in the gap under ellipsis? The answer to this question would provide invaluable information not only to the field of theoretical linguistics and language acquisition, but also to that of philosophy, cognitive science, information theory, natural language processing and machine learning, as it would help explain what the human brain considers critical for successful and efficient information transfer. In examining instances of successful information conveyance despite various inconsistencies in the elliptical signal, we would also gain information on the number and type of inconsistencies the human brain is willing to overlook in order to achieve successful communication. Finally, given one of the markers of cognitive decline in elderly populations and neurodegenerative disorders is problematic information identification and retainment, understanding how ellipsis works as a phenomenon and how such populations comprehend and produce it could provide us with an additional tool to help identify the extent and progression of such degradation patterns.

Despite ellipsis featuring prominently in syntactic and psycholinguistic literature over the past five decades, there is still great debate as to exactly what information is relevant to its success and how this information is conveyed. Two main theoretical camps exist: one argues the *e*-site contains full syntactic structure that conveys meaning without being pronounced; and the second that meaning is retrieved without such structure being present at the *e*-site. Deciding between these two views has been complicated by data behaving both consistently and inconsistently with unpronounced structure at the *e*-site. This thesis provides valuable evidence towards adjudicating between these two views and addresses two key gaps in the ellipsis literature. Firstly, it provides novel, adequately powered and controlled web-based cross-linguistic acceptability datasets that speak to the debate on structure within the *e*-site. In extension, a methodological guide for adapting and implementing similar experiments in other languages is provided. Secondly, this thesis offers a cohesive, cross-linguistically generalisable interdisciplinary account of ellipsis. Drawing on both the fields of theoretical syntax and information theory, it argues that there is structure at the site of ellipsis but that our imperfect memory system acts as a noisy channel allowing constrained structural infidelities according to precise informational criteria. Collectively, this work presents the most extensive interdisciplinary exploration of ellipsis to date.

Acknowledgements

To list all the people and events that have shaped the course of my PhD and the writing of this thesis would be too much for any reader to handle. Ironically, however, the greatest acknowledgements, those that should go without saying, are also those that are too important to leave out.

First of all, I would like to thank my amazing parents. There are no words to express how grateful I am for the love and support you have given me without question in every step I have taken throughout my life. Our daily conversations have diminished any obstacle I have ever faced and I know I would not have made it this far without them. This thesis is dedicated to you.

Thank you, of course, to my supervisors, Andrea and Klaus. Your input has been vital to all the ideas I have had and the experiments I have run from MSc to PhD. To Andrea; thank you for making me an experimentalist who will not accept any result without an accompanying graph; but also for all our conversations on any subject, for your constructive but kind criticism and your belief in me. To Klaus; thank you for introducing me to the subject of sluicing and convincing me to undertake this PhD. Thank you also for your enthusiasm in all my ideas and for every syntax discussion that has helped push my boundaries. You are polar opposites of each other and it is your combined support and guidance that has made me the academic I am today.

Thank you to Jason Merchant for giving me the opportunity to visit the University of Chicago and - together with Anastasia - making me feel so welcome; for all the in-depth discussions, exceptional insight and feedback on my ideas. To Masaya Yoshida, thank you for welcoming me into your lab, for being an overseas mentor and for all the valuable discussions in both person and via Google Hangouts.

Thank you to Elena, for every frivolous and serious discussion. You have proven to be one of my dearest friends and I know I can talk to you about anything.

And last, but the opposite of least, to Toby. You are the calm to my storm and I could not imagine my life without you.

Contents

1 Sluicing and its theories **15**

1.1 Introduction . . . 15

1.1.1 What is sluicing? . . . 15

1.1.2 What are the main features of sluicing? . . . 16

1.2 Theories of Sluicing . . . 19

1.2.1 Most Prominent Syntactic Theories of Sluicing . . . 20

1.2.2 PF-Deletion and the Amelioration of Islands, a First Account . . . 22

1.2.3 LF-Copying and Sprouting . . . 24

1.2.4 PF-Deletion Revisited . . . 28

1.2.5 Pointer or Anaphor Account . . . 32

1.3 P-Stranding under Sluicing in Non-P-Stranding Languages . . . 37

1.3.1 Two Classes of Explanation . . . 37

1.3.2 PF-Deletion in spite of P-Stranding . . . 38

1.3.3 P-Stranding therefore No PF-Deletion . . . 49

1.3.4 Conclusion and the Cases of Greek and German . . . 53

2 Cross-linguistic Acceptability Judgement Data **57**

2.1 Experiment 1: P-Stranding under Regular Sluicing . . . 57

2.1.1 Experimental Outline, Aim and Predictions . . . 57

2.1.2 Methods . . . 64

2.1.3 Results . . . 73

2.1.4 Discussion . . . 80

2.2 Experiment 2: P-Stranding out of Islands under Sluicing . . . 81

2.2.1 Experimental Outline, Aim and Predictions . . . 81

2.2.2 Methods . . . 89

2.2.3 Materials . . . 89

2.2.4 Results . . . 91

2.2.5 Discussion . . . 97

2.3 Experiment 3: P-Stranding under Contrast Sluicing . . . 99

2.3.1 Experimental Outline, Aim and Predictions . . . 99
2.3.2 Methods . . . 101
2.3.3 Materials . . . 101
2.3.4 Results . . . 104
2.3.5 Discussion . . . 107
2.4 General Discussion for all Greek Acceptability Judgement Studies . . . 107
2.5 Experiment 4: P-Stranding under Sluicing in German . . . 113
2.5.1 Experimental Outline, Aim and Predictions . . . 113
2.5.2 Methods . . . 116
2.5.3 Results . . . 122
2.5.4 Discussion . . . 135
2.6 General Discussion of both Greek and German data . . . 139
2.6.1 P-stranding via an Alternative Source . . . 140
2.6.2 P-Stranding as a Grammatical Illusion . . . 146
2.6.3 Is Stranding the P the same as Overtly Dropping it? . . . 147
2.6.4 Greek and German are *not* ‘PSG-exceptions’ . . . 149

3 A Noisy Channel Model of Sluicing 151

3.1 Introduction . . . 151
3.2 Anticipation, Bayesian Estimation & Noise . . . 152
3.2.1 The What; Anticipation and/or Predictability . . . 152
3.2.2 The How; Frequency, Probability & Surprisal . . . 155
3.2.3 The Why; Uncertain Input & the Noisy Channel model . . . 161
3.3 Sluicing as a Noisy Channel . . . 167
3.3.1 ‘P-Stranding’ under Sluicing . . . 167
3.3.2 Case-Mismatching under Sluicing . . . 178
3.3.3 Island amelioration under Sluicing . . . 185
3.4 Conclusion . . . 189

4 Processing Sluicing 194

4.1 Introduction . . . 194
4.2 Background: Yoshida, Dickey & Sturt, 2013 . . . 196
4.2.1 Study Design & Results . . . 196
4.2.2 Objections . . . 204
4.2.3 Conclusion . . . 207
4.3 Self-Paced Reading Replication in Modern Greek; . . . 208
4.3.1 Experimental Outline, Aim and Predictions . . . 208
4.3.2 Methods . . . 209
4.3.3 Results . . . 216

4.3.4 Discussion . . . 222
4.4 Simplified English Self-Paced Reading Replication . . . 223
4.4.1 Experimental Outline, Aim and Predictions . . . 223
4.4.2 Methods . . . 228
4.4.3 Results . . . 231
4.4.4 Discussion . . . 237
4.5 Sentence Completion Study . . . 240
4.5.1 Experimental Outline, Aim and Predictions . . . 240
4.5.2 Methods . . . 244
4.5.3 Results . . . 245
4.5.4 Discussion . . . 249
4.6 Overall Discussion . . . 252

5 Conclusions and Future Directions 257

5.1 Summary . . . 257
5.2 Limitations and Future Directions . . . 273
5.3 Concluding Contributions . . . 278

Appendices 280

A Appendix: Chapter 1 281

A.1 Nykiel (2013) Polish Experiment Design . . . 281

B Appendix: Chapter 2 283

B.1 Experiment 1: P-Stranding under Regular Sluicing (Greek) . . . 283
B.1.1 Items . . . 283
B.2 Experiment 2: P-Stranding under Sluicing in the context of Islands; Stimuli . 306
B.2.1 Experimental Items . . . 306
B.3 Experiment 3: P-Stranding under Contrast Sluicing; Stimuli . . . 309
B.3.1 Experimental Items . . . 309
B.4 Further Information on Dative Alternation in Modern Greek . . . 312
B.5 Experiment 4: P-Stranding under Regular Sluicing (German) . . . 314
B.5.1 Items . . . 314
B.5.2 Additional Results Tables . . . 329

C Appendix: Chapter 4 333

C.1 Greek SPR Study . . . 333
C.1.1 Residual RT Mean and SEM Results per Region per Condition . . . 333
C.1.2 Raw RT Mean and SEM Results per Region per Condition . . . 335
C.1.3 Order Effect Analysis Results . . . 337

C.1.4 Greek SPR Stimuli 340

C.2 Simplified English SPR Study 362

C.2.1 Residual RT Mean and SEM Results per Region per Condition 362

C.2.2 Raw RT Mean and SEM Results per Region per Condition 365

C.2.3 English SPR Stimuli 367

C.2.4 English SC Stimuli 379

List of Tables

1.1 Nykiel (2013) Exp. 1 Overview . . . 53
1.2 PSG-defying Languages, alphabetically . . . 56

2.1 Experiment 1 Conditions and Examples . . . 59
2.2 Nykiel (2013) Pointer Account Acceptability Predictions . . . 63
2.3 Experiment 1 Significance Predictions . . . 63
2.4 Experiment 1 Conditions and Examples . . . 65
2.5 Experiment 1 Stimuli Organisation . . . 70
2.6 Experiment 1: Acceptability Ratings Numerical Summary . . . 77
2.7 Experiment 1: Main Effects & Interactions . . . 78
2.8 Experiment 2 Conditions and Examples . . . 84
2.9 PF-Deletion Acceptability Predictions for P-Stranding, Sluicing & Islands . . . 85
2.10 PF-Deletion Acceptability Predictions for P-Stranding, Sluicing & Islands, with PPs as PF-Islands . . . 86
2.11 LF-Copying Acceptability Predictions for P-Stranding, Sluicing & Islands . . . 86
2.12 Pointer Account (Nykiel, 2013) Acceptability Predictions for P-Stranding, Sluicing & Islands . . . 88
2.13 Experiment 2 Significance Predictions . . . 88
2.14 Experiment 2 Conditions and Examples (Repeated) . . . 90
2.15 Experiment 2: Acceptability Ratings Numerical Summary . . . 93
2.16 Experiment 2: Main Effects & Interactions . . . 93
2.17 Experiment 3 Conditions and Examples . . . 102
2.18 Experiment 3: Acceptability Ratings Numerical Summary . . . 105
2.19 Experiment 3: Main Effects & Interactions . . . 106
2.20 Greek Dative Alternation pattern by dialect . . . 111
2.21 Experiment 4 Predictions . . . 116
2.22 Experiment 4 Conditions and Examples . . . 117
2.23 Experiment 4 Stimuli Organisation . . . 121
2.24 Experiment 4: Acceptability Ratings Numerical Summary . . . 126

2.25 Experiment 4: Main Effects & Interactions 127
2.26 Thermometer Items: Acceptability Results Summary 129
2.27 Greek & German; Acceptability Ratings Numerical Summary 137
2.28 Greek & German; Main Effects & Interactions 138

3.1 Cross-linguistic P-Stranding Predictions for Non-P-Stranding Languages . . . 173

4.1 Mean Acceptability Ratings per Condition 202
4.2 Greek SPR Conditions and Examples 210
4.3 Greek SPR Region Collapse 217
4.4 Greek SPR Analysis Results Summary by Region 220
4.5 Greek SPR Analysis Results Summary by Region (Continued) 220
4.6 Simplified English SPR Conditions 226
4.7 Simplified English SPR Conditions (annotated) 228
4.8 English SPR Analysis Results Summary by Region (Main Effects) 235
4.9 English SPR Analysis Results Summary by Region (Interactions) 235
4.10 SC Study Conditions and Examples 242
4.11 SC Results; Continuation Percentages by Condition 247
4.12 SC Chi-Square Results 250

5.1 Cross-linguistic P-Stranding Predictions for Non-P-Stranding Languages . . . 268
5.2 Follow-up Noisy Channel Study Conditions 276

B.1 Greek & German Three-Way Interactions: Case-Matching*P-Stranding*Sluicing330
B.2 Greek & German Three-Way Interactions: Language*Case-Matching*Sluicing 330
B.3 Greek & German Three-Way Interactions: Language*P-Stranding*Case-Matching331
B.4 Greek & German Three-Way Interactions: Language*P-Stranding*Sluicing . 331
B.5 Greek & German Four-Way Interaction 332

C.1 Greek SPR RT Means and SEM per Region per Condition 333
C.2 Greek SPR Linear Regression Analysis Results 337
C.3 English SPR RT Means and SEM per Region per Condition 362

List of Figures

1.1 Terminology Used . . . 16
1.2 Phillips and Parker (2014) Ellipsis Decision Tree . . . 21
1.3 Account Classification . . . 22

2.1 Frequency effect on reaction times . . . 72
2.2 Experiment 1: Mean Acceptability Ratings Barplot . . . 77
2.3 Experiment 1: Fillers Mean Acceptability Ratings Barplot . . . 78
2.4 Experiment 2: Mean Acceptability Ratings Barplot . . . 92
2.5 Experiments 1 and 2: Mean Acceptability Ratings Barplot . . . 95
2.6 Experiment 3: Mean Acceptability Ratings Barplots . . . 105
2.7 All Greek Acceptability Results . . . 108
2.8 Experiment 4: Acceptability Ratings Barplot . . . 126
2.9 Experiment 4: Fillers Mean Acceptability Ratings Barplot . . . 127
2.10 Thermometer Items: Acceptability Ratings Barplots . . . 129
2.11 Thermometer Items: Acceptability Ratings Scatterplot per Subject . . . 130
2.12 Greek & German: Acceptability Ratings . . . 136

3.1 Shannon 1948 General Communication Model . . . 162
3.2 Shannon 1948 Annotated General Communication Model . . . 164
3.3 Greek & German Case-Matching Results . . . 175
3.4 AlShaalan and Abels (2019) Mean Acceptability Ratings Barplot . . . 177
3.5 Greek & German Case-Mismatching Results . . . 181
3.6 Experiment 4: Acceptability Ratings Barplot . . . 184
3.7 Experiment 2: Mean Acceptability Ratings Barplot . . . 187
3.8 Experiments 1 and 2: Mean Acceptability Ratings Barplot . . . 188

4.1 All Yoshida, Dickey, and Sturt (2013) SPR results . . . 200
4.2 Interaction Summary . . . 201
4.4 Greek SPR Log Transformed Results; All Conditions . . . 219

4.6 English SPR Results; All Conditions 234
4.7 PP Type Means with SEM for C2(*and*); Pied-Piped Conditions only 237
4.8 Sentence Completion Results; Percentages by Condition 247

Chapter 1

Sluicing and its theories

1.1 Introduction

1.1.1 What is sluicing?

Despite sound pervading most aspects of language - and indeed of life - and being a basic medium for conveying meaning in everyday interactions, there exists a phenomenon which represents one of the most profound breakdowns of this sound-meaning coupling outside of sign language. This phenomenon, whereby meaning is somehow successfully conveyed through silence, i.e. in the absence of overt linguistic expression, is known as *ellipsis* - from the Ancient Greek 'to miss' ($\lambda\epsilon\acute{\iota}\pi\omega$) - and has confounded linguists the world over for decades. Ellipsis appears in various forms, each targeting a different sentence constituent for omission, such as NPs (2) or VPs (3). What all forms of ellipsis have in common, however, is their apparent ease of inference or interpretation from the surrounding context, leading most to speculate that redundancy must be a necessary, if not sufficient, prerequisite of this phenomenon. As such, ellipsis can be argued to be a type of linguistic economy mechanism, a part of the brain's - and indeed nature's - love of conserving energy whenever possible.

(2) Mary ate two slices of cake, whereas her brother ate five [~~slices of cake~~].

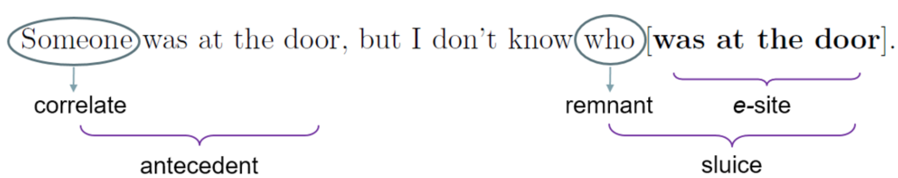
(3) Mary's parents loved her brother and Mary begrudgingly did [~~love her brother~~] too.

(4) Mary was sick of someone eating all her cake, and we all knew who [~~she was sick of t_i eating all her cake~~].

The form of ellipsis which will concern this thesis in particular is that of *sluicing*, where an interrogative clause is apparently reduced to its introductory *wh*-element (4). What makes this type of ellipsis particularly interesting are its two most (in)famous characteristics, pointed out by John Robert 'Haj' Ross in 1969: the fact that sluicing appears to both comply with but also defy overt syntactic rules, thus presenting a remarkable conundrum for syntacticians and semanticists for over half a century. To introduce these characteristics, I will be adopting the terminology commonly used in the sluicing literature and shown in figure (1.1) for the simple example in (5). Specifically, the *wh*-element which appears to be left behind after deletion is known as the sluice *remnant*; the content which is presumed to be deleted or *elided* is considered to be inside the *site of ellipsis* or *e*-site; together, remnant and *e*-site constitute the *sluice*. The sentence upon which the sluice appears to be based and from which we can infer its content is known as the sluice's *antecedent* and the indefinite contained therein, which the remnant refers to and appears to be co-indexed with, is known as its *correlate*.

(5) Someone was at the door, but I don't know who.

Figure 1.1: Terminology Used



The rest of this chapter will be concerned with introducing the key features that make sluicing such an intriguing subject of study, along with the various theories which have been proposed to capture its behaviour.

1.1.2 What are the main features of sluicing?

The first key feature of sluicing is how it appears at first glance to adhere to the rules of overt syntax. That is to say, the *wh*-element exhibits the vast majority of characteristics one would expect of it if indeed it was the remnant of an otherwise regular interrogative CP, i.e. if it had arrived at this position via regular *wh*-movement. Firstly, in overtly case-marking languages, it bears the specific morphological case one would predict if the content within the *e*-site had been overtly expressed with a clause identical to the antecedent - save for the *wh*-phrase (e.g. (6), where the verbs *schmeicheln* and *loben* pattern with Dative and Accusative respectively). Secondly, if within the antecedent the sluice correlate appears as the complement of a preposition (P), then it seems as though the sluice remnant can only appear without this P if the language is an overtly P-stranding language (e.g. (7) vs. (8)). These two features are collectively known as *connectivity effects* and form the basis of some of the most influential theories of sluicing which we will return to shortly.

(6) *Case-Matching: German*[1]

a. Er will jemandem schmeicheln, aber sie wissen nicht, *wer
*he wants someone.*DAT *flatter but they know not who.*NOM
/*wen /wem [~~er schmeicheln will~~].
*who.*ACC *who.*DAT *he flatter wants*
'He wants to flatter someone, but they don't know who.'

b. Er will jemanden loben, aber sie wissen nicht, *wer /wen
*he wants someone.ACC. praise but they know not who.*NOM *who.*ACC
/*wem [~~er loben will~~].
*who.*DAT *he praise wants*
'He wants to flatter someone, but they don't know who.'

(7) *Preposition-Stranding: English*

a. Who did Mary dance with?

b. Mary danced with someone, but I do not remember (with) whom.

(8) *Preposition-Stranding: German*

a. * Wem hat Maria mit getanzt?
*who.*DAT *has Maria with danced*
'Who did Maria dance with?

b. Maria hat mit jemandem getanzt, aber wir wissen nicht *(mit) wem.
*Mary has with someone.*DAT *danced but we know not with who.*DAT
'Mary danced with someone, but we do not know (with) whom.'

Simultaneously, however, this same *wh*-element appears to defy other overt syntactic rules, consistently escaping what would traditionally constitute syntactic islands should this CP have been overt (9), i.e. phrases from within which elements are not permitted to move out (Ross, 1967). Although not all islands appear to be ameliorated through sluicing (see also Chapter 2.3), and there is much debate as to whether these are true instances of island escape or not (see, e.g. Abels (2011); M. Barros, Elliott, and Thoms (2014); Gribanova (2013), to mention but a few), this feature is referred to as an *anti-connectivity effect*. Together, these two sets of effects have shaped a multitude of theoretical approaches to sluicing based on syntax, semantics, phonology and/or sentence processing.

[1]Example taken from Merchant, 2001: p. 42, his (13) & (14), in turn from Ross (1969)

(9) *Island Ameliorations under Sluicing*[2]

a. Relative Clause

i. * They wanted to hire someone who speaks a Balkan language, but I don't remember [which Balkan language]$_i$ they wanted to hire someone [$_{\text{CP}}$who speaks t_i].

ii. They wanted to hire someone who speaks a Balkan language, but I don't remember which.

b. Complex NP

i. * The rumour that the government is planning to appoint someone as new home secretary has just been confirmed, but I cannot remember who$_i$ [$_{\text{NP}}$the rumour that the government is planning to appoint t_i as new home secretary] has just been confirmed.

ii. The rumour that the government is planning to appoint someone as new home secretary has just been confirmed, but I cannot remember who.

c. Adjunct

i. * The detective had to leave the building before interviewing someone, but we did not know who$_i$ he had to leave the building [$_{\text{AdvP}}$before interviewing t_i].

ii. The detective had to leave the building before interviewing someone, but we did not know who.

Both of these types of effects have been summarised in the seminal thesis and subsequent book by Jason Merchant 2001 in the form of (10) - (12), which we will return to frequently throughout this thesis.

(10) **Connectivity Effect: Case-Matching**
(a.k.a. Form Identity Generalisation I (Merchant, 2001, p. 91))
The sluiced *wh*-phrase must bear the case that its correlate bears.

(11) **Connectivity Effect: P-Stranding**
(a.k.a. Form Identity Generalisation II (Merchant, 2001, p. 92))
A language *L* will allow preposition stranding under sluicing iff *L* allows preposition stranding under regular *wh*-movement.

[2]Only a few examples of possible island ameliorations are given here, with other ameliorations including Coordinate Structure, Sentential Subject, Left Branch Condition, COMP-trace effect and Embedded Question violations.

(12) **Island Amelioration**
Sluicing with an overt correlate does not respect syntactic islands[3].

The main focus of this thesis shall be on one of these connectivity effects, that of Preposition-Stranding or P-Stranding (11), also known as the P-Stranding Generalisation (PSG), as well as secondarily the Case-Matching connectivity effect or Case-Matching Generalisation (CMG). After introducing the various theories of sluicing, we will discuss here the existing cross-linguistic literature on the subject and how it has provided varied and sometimes contradictory evidence of this behaviour, leaving a convoluted and confusing scene for researchers to decipher. Specifically, even though the majority of languages have been argued to respect these two generalisations, we will be looking more closely at instances where the PSG, in particular, appears to be defied, with certain overtly non-P-stranding languages appearing to allow remnants to be stranded from their Ps under sluicing. One major issue of this existing literature, however, is that there is no clear, cross-linguistically applicable explanation proposed for this behaviour thus far. A further major problem which plagues it is the lack of properly powered experiments documenting the true appearance of this phenomenon in naïve native speakers. The purpose of this thesis shall, therefore, be to provide such novel experimental evidence documenting similar defiant behaviour in two languages which have previously been proposed to be some of the most robust examples of the PSG, namely Modern Greek and German. From there, I will move on to propose a cross-linguistically generalisable explanation of this behaviour rooted in both theoretical syntax and sentence processing.

Before we dive into this phenomenon of P-stranding under sluicing, however, let us first take a look at sluicing more generally and at the main approaches which have been proposed to account for its multifaceted behaviour, along with the advantages and drawbacks of each.

1.2 Theories of Sluicing

Thus far in the literature, the examination of ellipsis in general, and sluicing more specifically, can be easily split into roughly three levels of analysis or examination, following the very useful representation given in Phillips and Parker (2014) and repeated here in fig. 1.2. On the first level of analysis as depicted, there is the literature targetting the nature of the antecedent and its relationship to the sluice; this level of analysis attempts to find the key common features of antecedent and sluice which will allow the licensing of sluicing, in the sense of whether there must be some form of stricter or more flexible semantic, discourse

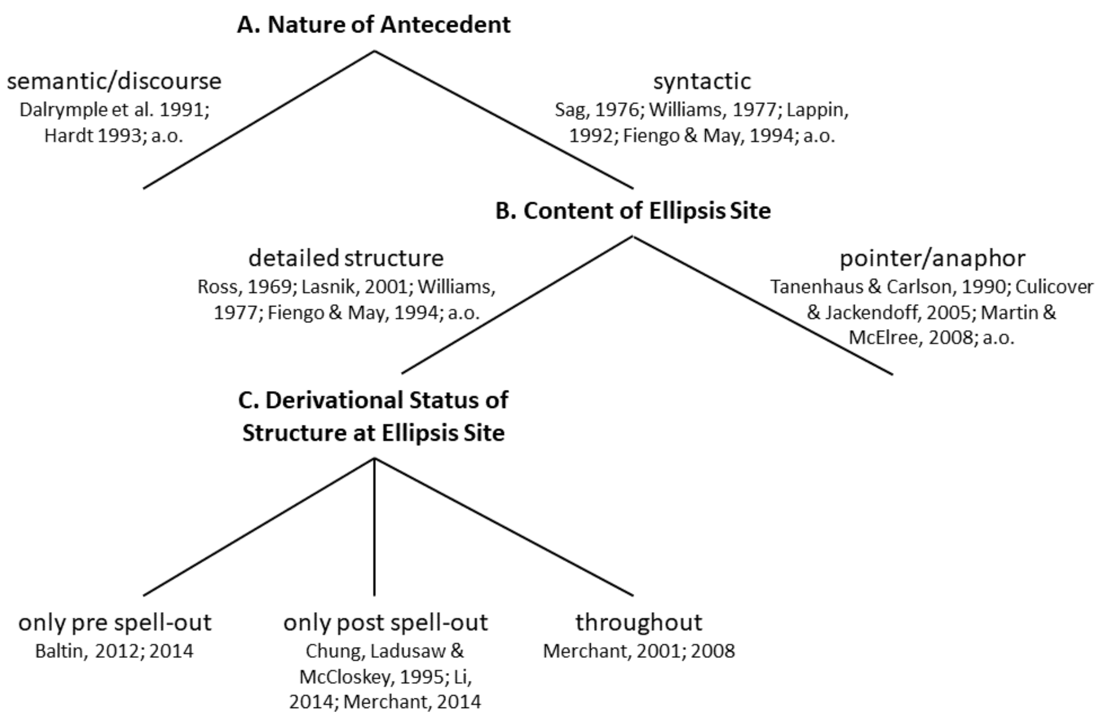
[3]Here I use the term 'syntactic' to mean what Merchant (2001) termed 'strong' islands undone by PF-Deletion, i.e. Left-Branch extraction, COMP-trace effects, topicalisations and subject extraction, with extraction out of conjuncts, complex NPs and adjuncts considered only superficially ameliorated under PF-Deletion; selective islands, in contrast, are considered semantic/pragmatic in nature (contra Manzini (1998); Rizzi (1990)) and hence not relevant at the PF level.

and/or syntactic parallelism restriction between said antecedent and *e*-site. The second level of examination deals with the content of the *e*-site itself; the debate here refers to whether the site of ellipsis itself is thought to contain full, detailed syntactic structure which is simply not pronounced (for whatever reason), or whether instead it contains not structure, but some form of processing mechanism, such as a pointer or anaphor, either of which might serve to guide the parser back to the antecedent in order to make full semantic interpretation possible. Finally, if we accept that there *is* some form of structure at the *e*-site and not a pointer, then the third level of analysis deals with the question of what this structure's derivational status is, i.e. exactly when it appears and/or disappears from this site with respect to the different levels of representation as defined by traditional generative syntax. Given both spatial and temporal constraints, in this document we will not deal with all three of these levels, directing the reader instead to other, much more thoroughly analysed and argued treatises wherever possible; we will instead argue simply for there being detailed structure at the *e*-site, without strictly committing to either its derivational status or its precise relation to the antecedent. This is because the topic of interest here lies not specifically in the theoretical licensing or syntactic derivation of *e*-site material, as has concerned much of the more well-known literature, but rather in how the parser more practically *deals* with ellipsis. Specifically, we will examine how the parser is equipped to overcome what have previously been argued to be theoretically impossible situations, i.e., primarily, P-stranding under sluicing in overtly non-P-stranding languages, but also, secondarily, case mismatching under sluicing. Since we are only interested in these phenomena, we will narrow down our attention predominantly to level B, presenting now a brief overview of the most influential theories at this level of analysis and in interaction with levels A and C, and referring the reader to Lasnik (2001); Merchant (2001); Phillips and Parker (2014) for more extensive analyses and reviews.

1.2.1 Most Prominent Syntactic Theories of Sluicing

There are several syntactic theories on the derivation, and, by extension, licensing and identification, of sluicing, ranging from those based on more or less traditional syntactic and semantic theory to purely sentence processing ones. Here, focusing mainly on the second level of Figure 1.2 and its interaction with the third level, four main approaches to sluicing will be described, three of which have arguably shaped the field, at least as far as theoretical syntax is concerned. These three argue for detailed structure being present at *some* point within the *e*-site, whereas the last one argues for no structure at any point. In order to discuss their strengths and weaknesses, I will split them slightly differently into two groups, classifying them as 'movement' vs. 'non-movement' (Figure 1.3), based on whether they suppose the sluice remnant arrived at its surface position through regular *wh*-movement or not. As a result of this difference, the two groups are complementary with respect to

Figure 1.2: Phillips and Parker (2014) Ellipsis Decision Tree



whether they naturally capture connectivity or anti-connectivity effects.

Figure 1.3: Account Classification

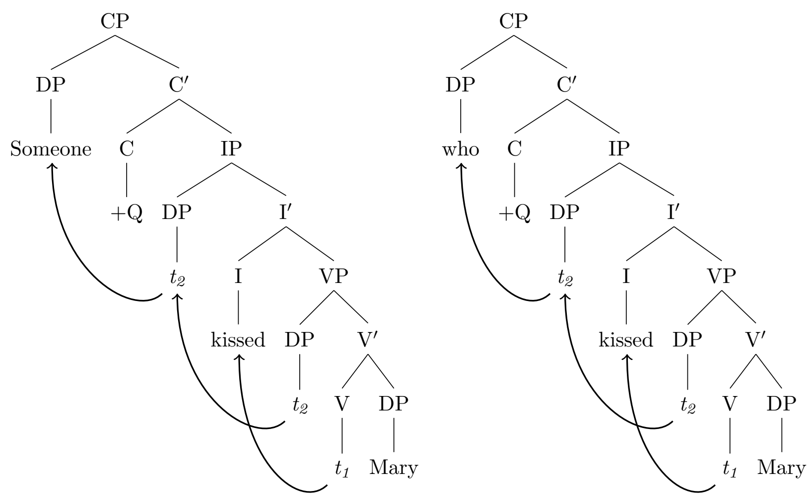
Movement accounts, on the one hand, as the name implies, consider the *wh*-remnant's surface position to be the result of regular syntactic transformations with the sluice once being a full interrogative CP whose IP has been deleted prior to spell-out. In other words, they posit detailed syntactic structure at the *e*-site pre spell-out which exists in a relationship of syntactic and/or semantic identity with the antecedent. This position simultaneously presents their biggest strength and weakness; by proposing the sluice was once an interrogative CP, with the remnant receiving case-marking as per usual and arriving at [Spec, CP] via regular *wh*-movement, both the CMG and PSG naturally follow, however island amelioration needs to be separately accounted for somehow. Non-movement accounts, on the other hand, argue that the sluice remnant was instead base-generated in its surface position without involving *wh*-movement. Out of these, one argues that the sluice's meaning is obtained via copying over the antecedent's IP post spell-out to form a full CP at LF, whilst the other simply that the *e*-site contains a pointer mechanism guiding the parser back to the antecedent to retrieve the sluice's meaning from there, with no structure ever existing in the *e*-site. By postulating no *wh*-movement in the *e*-site, these last two accounts easily account for islands not presenting an issue for sluicing; however, they are unable to capture the two connectivity effects without also positing additional constraints on the form of the remnant.

1.2.2 PF-Deletion and the Amelioration of Islands, a First Account

In the first theoretical attempt to explain connectivity and anti-connectivity effects, Ross (1969) instigated the debate on the existence of syntax at the *e*-site with his *PF-Deletion Hypothesis*. According to this approach, a full interrogative CP is base-generated at the *e*-site under strict syntactic identity with preceding material. This CP is subject to regular transformations, however at some point, prior to its phonological realisation, this CP's IP is deleted (e.g. (13) would have (14) as its source - with only the relevant matching parts of the derivation shown here).

(13) Someone kissed Mary, but I don't remember who.

(14)



As expected, such a deletion hypothesis can easily account for the uniform behaviour of case-matching and P-stranding as phenomena both under sluicing and overtly. With respect to island ameliorations, however, in order to overcome this stumbling block, Ross proposed two further additions to his main theory; firstly, that a sentence's acceptability is calculated over the entirety of the derivation by some form of global rule; and secondly, that only island-forming nodes appearing at surface structure are considered truly unacceptable (Ross, 1969, p. 277). In this way, sluicing would inherently ameliorate all islands, as these illegal, island-forming nodes are never phonologically realised and would, thus, never be seen by a globally calculating rule. This generalisation was later reiterated in Lakoff (1972) and Baker and Brame (1972) such that the crossing of an island-node assigns to it an ungrammatical feature, with the subsequent deletion of this feature under sluicing resulting in grammaticality (see also Fox and Lasnik (2003); Merchant (2008) for more updated versions). However one frames this generalisation, though, it presents with two major overgeneralisation drawbacks. Firstly, it would also predict islands to be ameliorated in other instances of ellipsis which contain such an ungrammatical feature, such as in VP-Ellipsis; and secondly, it would predict *all* islands to be ameliorated under sluicing. Neither of these predictions is borne out (e.g. VP-ellipsis in (15) vs. equivalent sluice in (9a), repeated here), as we will also see below when discussing a particular case of sluicing known as 'sprouting'.

(15) * The board wanted to hire someone who speaks a Balkan language, but I don't remember which (Balkan language) the chairman did [~~want to hire someone who speaks~~].

(9a) They wanted to hire someone who speaks a Balkan language, but I don't remember which (Balkan language) [~~they wanted to hire someone who speaks~~].

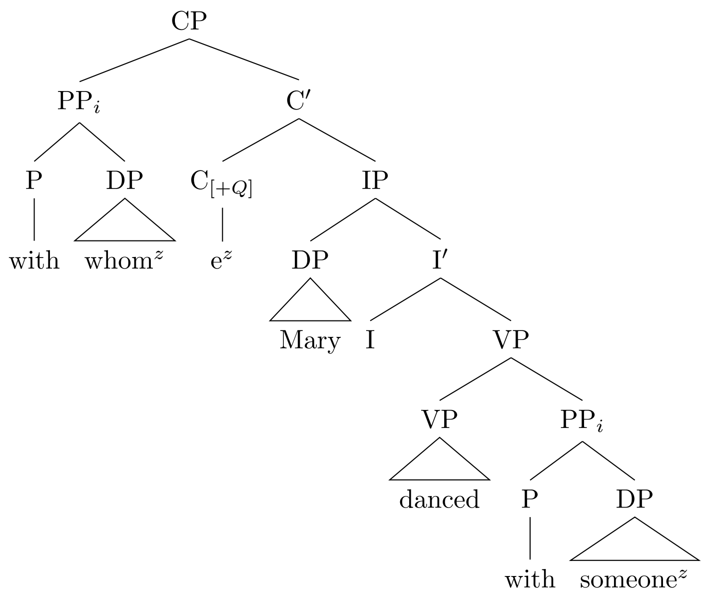
1.2.3 LF-Copying and Sprouting

An alternative approach was taken several decades later, in 1995, by Chung, Ladusaw, and McCloskey (henceforth CLM) with their theory of *LF-Copying*. This semantically-driven approach follows in the footsteps of Wasow (1972), Williams (1977) and Chao (1987). As opposed to PF-Deletion, LF-Copying posits no real structure at the *e*-site at the point of derivation, with the *wh*-phrase instead being base-generated alone in [Spec, CP] (16). Interpretation is achieved via a copying process of the antecedent IP (termed *recycling*) into the empty *e*-site later on, at LF, after the structure has been sent to PF for pronunciation[4]. The core of LF-Copying are the necessary syntactic and semantic binding conditions which accompany this recycling operation in order to reach the desired interpretation for the copied IP. Firstly, the base-generated *wh*-element must syntactically bind some available position within the copied IP. Secondly, this same element must also contain an indefinite (e.g. here, *someone*) which is available to be semantically bound by Q - the interrogative operator in C. This leads to a process they term *Merger*, during which the *wh*-phrase and copied available indefinite are both bound by the same operator and fuse semantically, with the resulting LF output inheriting their combined semantic properties. This is able to happen since, according to CLM, variables with unspecified referents are indistinguishable during interpretation and can, consequently, be unified (or *merged*). For instance, in (16), *with* can be merged with *with* on the basis of identity, whereas *whom* and *someone* are able to be merged[5], according to CLM, given they are indefinites without specified referents. As such, the grammaticality of sluicing is tied to the availability of an unbound variable (here the indefinite) in the copied IP.

[4]In using temporal terms such as 'later on', 'prior' etc., we are simply following syntactic practice and do not intend to convey that these operations occur in a temporally distinct manner. Indeed, in most of the parsing literature, these are not considered to be separate real-time processes, with the parser instead presumed to simultaneously compile all such levels in parallel (Fodor, Bever, Garrett, et al., 1974). See also Phillips and Parker (2014) and Phillips and Lewis (2013) for review.

[5]Indicated by cosuperscription.

(16)



Since this account is not based on *wh*-movement of the remnant, its island insensitivity is a natural consequence. This mechanism, furthermore, successfully captures a large number of other cases where sluicing is not permitted, such as when the potential antecedent is a referential expression ((17), their (28c)), a quantificational DP or pronoun ((18), their (30c)), or cases when it is within the scopal domain of negation[6] ((19), their (36c)), since none of these present an available unbound indefinite antecedent.

(17) * Because we suspected Joe had given it to Max, we then asked to whom/who to.

(18) * She's read most books, but we're not sure what/which.

(19) * No one signed any documents, but he's not sure what/which.

A particularly important contribution of CLM's paper, however, was to identify cases where the antecedent IP does not contain an overt correlate of the remnant (e.g. in (20), the antecedent does not contain an overt internal argument for *paint*, hence *what* has no overt correlate). These are termed instances of *sprouting*, as the recycled IP does not already contain an available position for the *wh*-element to bind (such as *someone* in (16)) and hence one must be created or *sprouted*. This position is constrained by the nature of the argument structure within which or adjoining to which it is created. Perhaps more importantly, however, sprouting seemingly differs from thus far encountered instances of sluicing, in that it appears to *obey* certain island constraints[7] (21), thus presenting a veritable crux for Ross's

[6]This is predicted following Heim's theory of indefinites, according to which, an indefinite either is and remains free or becomes bound within another operator's scopal domain, e.g. in this case the scope of negation.

[7]At least selective or 'weak' islands (Albert, 1993; Romero, 1998; Sauerland, 1996).

general island amelioration under sluicing position. The way that CLM incorporate this into their account is by considering sprouting to not be a case of Merger, as with regular sluicing, but rather a case of FormChain, with island constraints behaving as a restriction not on movement, but on $\overline{\text{A}}$-chain formation (Cinque, 1990).

(20) Mary wanted to paint, but it was not clear what.

(21) **Sprouting Islands**
Wh-Island; Subject Island; Complex NP Island
a. * Louie was thinking how to paint, but it was not clear what.
b. * That the guests would want to drink was likely, but we did not know what.
c. * They wanted to hire someone who drew well, but it was not clear what.

Of course, the great disadvantage of this theory, is that by forgoing *wh*-movement it does not inherently capture connectivity effects. In order to do so, CLM impose an additional, structural isomorphism condition between correlate and remnant, implemented via this LF-Copying system, where, in order for ellipsis to be licensed, the *wh*-element must match the case and category of its antecedent. Although this would account for the CMG, it still does not capture the second connectivity effect, the PSG, that was observed a few years later by Jason Merchant (2001). The only way to capture this would be if the structural isomorphism that CLM call for between remnant and correlate extends not only to case, but also to phrasal category, i.e. a PP correlate should necessitate a PP remnant. The issue here, of course, would be that such an extension would force even overtly P-stranding languages to always have P-pied-piped sluice remnants if the antecedent also contains P-pied-piping, i.e. (22) should be considered unacceptable in English. This is evidently an incorrect overgeneralisation.

(22) Mary danced with someone, but I can't remember who.

Another issue however, which deserves to be mentioned at this point, is that by forcing remnant and correlate to always bear the same case, certain, restricted instances of case-mismatching between the two which have since been reported in the literature, cannot be accounted for. Examples of such mismatching come from Greek (23 - 24) (Molimpakis, 2016; also this document, 2.4), but also Turkish (Ince, 2009, 2012), Mongolian, Korean, Uzbek, Japanese and Chamorro (Vicente, 2015)[8]. Icelandic also shows this type of remnant alternation in fragment answers (25, 26) (Wood et al., 2016, (41, 43)), which have been argued to

[8] Although the status of some of the latter as clear sluicing is questionable (Wood, Barros, & Sigurðsson, 2016)

operate similarly to sluicing (Griffiths and Lipták (2014); Merchant (2001, 2005a), though cf. Morgan (1973)). The key in these instances of case-mismatching is that the antecedent verb can assign two different cases to the same argument with the same meaning. When this is true, it appears possible to have the correlate appear in one of these cases and the remnant in the other. Although this type of alternation would make sense for an approach advocating the sluice remnant is the product of regular *wh*-movement with the *e*-site containing a deleted copy of the antecedent verb, it is a major issue for non-movement-based approaches, as we shall also see below.

(23) Dino to fakelo kapjou /se kapjon.
Give.1SG *Det*.M.ACC *folder*.M.ACC *someone*.M.GEN *to someone*.M.ACC
'I give someone the folder/ I give the folder to someone'

(24) O Markos edose to fakelo se kapjon
Det.M.NOM *Markos*.M.NOM *gave*.3SG *Det*.M.ACC *folder*.M.ACC *to someone*.M.ACC
chthes arja to vradi, ala de thimame tinos /se
yesterday late Det.N.ACC *night*.N.ACC *but* NEG *remember*.1SG *who*.M.GEN *to*
pjon.
who.M.ACC
'Marcus gave the folder to someone late last night, but I do not remember who.'

(25) Mig /Mér langar að fara.
me.ACC *me*.DAT *wants to go*
'I want to go.'

(26) A. Hverjum langar að fara?
who.DAT *wants to go*
'Who wants to go?'

B. *Eg /Mig /Mér!
I.NOM *me*.ACC *me*.DAT
'Me!'

It should be noted, at this point, that a similar approach to that of CLM is Levin's (1982) version of LF-Copying, implemented within the Lexical-Functional Grammar framework. This account also posits base-generation of the remnant with LF-Copying of appropriate context-sourced material, however it differs with respect to the copying of the *wh*-remnant's correlate. Specifically, the process involves copying of the antecedent IP and co-indexation of the *wh*-indefinite and correlate. This is followed by deletion of the copied over correlate, and copying of the *wh*-indefinite into the now-empty position it occupied. Although close to CLM's LF-Copying approach, there are a couple of important differences. There is no process of Merger here, i.e. the variable bound by the operator Q does not inherit the properties of both correlate and *wh*-indefinite. Instead, the *wh*-indefinite 'takes over the grammatical and thematic functions of its antecedent', agreeing with it 'in case and other features' (Levin, 1982, p. 365). The PSG, nevertheless, remains uncaptured, as do the instances of case-mismatching.

1.2.4 PF-Deletion Revisited

In his seminal 2001 work, Merchant revives Ross's 1969 PF-Deletion-based hypothesis of ellipsis. In order to overcome the limitations of the original, however, he improves upon it in two significant ways.

Firstly, he forgoes a purely syntactic or semantic account of the licensing and recovery of e-sites in general, instead combining the two. This is done by abandoning the structural isomorphism condition on the licensing of ellipsis, which had been key to most previous theories, along with all the problems that plagued it[9]. Instead, he builds on focus conditions initially put forward by Rooth (1992), Romero (1998) and Schwarzschild (1999). These attempt to define when something can be considered 'given', i.e. semantically inferable from surrounding material; however, whereas these previous definitions focused on a one-way entailment of Q from P, where Q can represent the elided material and P some part of an antecedent clause, Merchant's revised focus condition on ellipsis contains a two-way entailment of Q and P ((27), (28), his (42) and (62) respectively).

(27) **e-GIVENness**
'An expression E counts as e-GIVEN iff E has a salient antecedent A and, modulo $\exists-$type shifting,

1. A entails F-clo(E)[10], and
2. E entails F-clo(A)'

(28) **Focus condition on IP-ellipsis**
'An IP α can be deleted only if α is e-GIVEN.'

Merchant (2001, pp. 26, 31)

[9]Some of the more evident problems of requiring structurally identical correlates and remnants are an account of the phenomenon of sprouting, as well as that of 'vehicle change', as defined by Fiengo and May (1994) (see further down for more details).

[10]Merchant (2001) defines the F-closure of α, written F-clo(α), as 'the result of replacing F-marked parts of α with $\exists-$bound variables of the appropriate type (modulo $\exists-$type shifting)' (Merchant, 2001, (8)) and $\exists-$type shifting as a 'type-shifting operation that raises expressions to type <t> and existentially binds unfilled arguments.' (Merchant, 2001, p. 14, footnote 3).

This two-way entailment goes above and beyond the original one-way version, in that it can capture subtle cases of non-mutual entailment, such as (29) having (30a) as a source, rather than (30b).

(29) Louise handed a glass of wine to Sarah, but I can't remember who else.

(30) a. Louise handed a glass of wine to Sarah, but I can't remember who else ~~Louise handed a glass of wine to~~.

b. * Louise handed a glass of wine to Sarah, but I can't remember who else ~~Louise served~~.

It also goes beyond sluicing in successfully handling what otherwise constitute bizarre ellipsis cases under any form of antecedent-copying approach based on a structural isomorphism condition. These include cases such as 'vehicle change' in VP-Ellipsis (Fiengo & May, 1994), where R-expressions apparently license the deletion of pronouns (e.g. (31a) and (32a) having (31b) and (32b) as sources, as opposed to (31c) and (32c), respectively.

(31) a. The teachers failed John$_i$, even though he$_i$ was hoping they wouldn't.

b. The teachers failed John$_i$, even though he$_i$ was hoping they wouldn't ~~fail him$_i$~~.

c. * The teachers failed John$_i$, even though he$_i$ was hoping they wouldn't ~~fail John$_i$~~.

(32) a. The teachers failed [one of the class's worst students]$_i$, even though he$_i$ was hoping they wouldn't.

b. The teachers failed [one of the class's worst students]$_i$, even though he$_i$ was hoping they wouldn't ~~fail him$_i$~~.

c. * The teachers failed [one of the class's worst students]$_i$, even though he$_i$ was hoping they wouldn't ~~fail [one of the class's worst students]$_i$~~.

One further, important aspect of such a PF-Deletion Hypothesis based on mutual semantic entailment is that it can theoretically capture those restricted instances of case-mismatching mentioned in the previous section, where the correlate and remnant may appear in a different case as long as a) both cases are licensed by the same antecedent verb and b) their meaning is identical. The only issue here is that the original Case-matching Generalisation states that remnant and correlate must appear in the same case. An important amendment to the PF-Deletion Hypothesis proposed in Abels (2016), however, overcomes this limitation. This amendment is known as the *Fit Condition* (33).

(33) **Fit Condition**

'Modulo agreement in the antecedent and wh-movement, replacing the correlate by the remnant in the antecedent must lead to a syntactically well-formed structure with the right meaning or – for sprouting – adding the correlate into the antecedent and making no further changes must lead to a syntactically well-formed structure with the intended thematic interpretation.'

Abels (2016, p. 9, (18))

This theoretical amendment essentially requires the remnant to appear in a case licensed by the antecedent verb as long as the meaning associated with these alternative case-marking instances remains identical; as such it can easily capture these examples of case alternation. It also successfully accounts for a number of otherwise puzzling situations for a PF-Deletion approach based purely on semantic identity, such as why cases along the lines of (34) are not interchangeable, as opposed to (35).

(34) a. Andrea drew a portrait, but I can't remember of who.
b. * Andrea drew a portrait, but I can't remember who.

(35) a. Andrea drew a portrait of someone, but I can't remember of who.
b. Andrea drew a portrait of someone, but I can't remember who.

In (34), both the main verb's argument structures should be semantically equivalent, with the preposition *of* being vacuous from a purely semantic point of view. Hence there does not appear to be a clear reason for only one of the two argument structures to be recoverable if we are to posit no clear syntactic identity condition between antecedent and remnant. The Fit Condition, by providing a small additional requirement to the semantic identity condition can however explain both these cases, by simply requiring that the remnant be able to 'slot' into the correlate's position non-modulo further changes. This condition brings to the fore the importance of the sluice remnant bearing not the case necessarily of its correlate, but rather the case required by the antecedent verb[11].

What I would like to focus mostly on, however, is Merchant's contribution of the PSG ((11), repeated here in (36)) to the form identity generalisations, which had up until this point consisted only of Case-Matching (10).

(36) **Form Identity Generalisation II**

A language *L* will allow preposition stranding under sluicing iff *L* allows preposition

[11] This amendment also addressed arguments that the judgements provided in Merchant (2001) could also be compatible with sluicing having *wh*-movement as a primary source, but a cleft or copular source as a secondary, 'last resort' source when *wh*-movement is not available (see Van Craenenbroeck (2010a) in particular for more details on this, as well as Chapters 2 and 3 here).

stranding under *wh*-movement.

Specifically, these two generalisations are proposed to form one class of behaviour under sluicing, matching their overt counterparts in conforming to regular transformational rules, with islands (for the most part) forming a separate behaviour class, one which does not obey these rules.

Finally, it should also be mentioned here that alongside the concept of PF-islands, there also exists an alternative approach to island amelioration under sluicing still in accordance with a PF-Deletion account. Specifically, given a lack of syntactic or structural isomorphism requirement between antecedent and sluice, this approach advances the concept that - at least some - superficially island ameliorating sluices are in fact derived of shorter, non-island-containing sources (e.g. (37) or (38)). These alternative sources are mostly considered to be copular or cleft sources, particularly for overtly non-case-marking languages or languages where there is case syncretism between the remnant case and a copular pivot case (Abels, 2017; Abels & Thoms, 2014; M. Barros et al., 2014; Erteschik-Shir, 1977; Fukaya, 2007, 2012; Gribanova, 2013; Marušič & Žaucer, 2013; Pollmann, 1975; Vicente, 2018). This approach is favoured by some, as it can account for the fact that not all islands are ameliorated under sluicing, with this distribution of acceptable and unacceptable island 'amelioration' appearing to mostly coincide with the availability or not of such alternative, short sources. Alternatively, following a similar logic, it is also proposed that such sources may not contain *wh*-movement per se, but rather scrambling or focus movement (Ince, 2009, 2012; Van Craenenbroeck & Lipták, 2006, 2013). Although these alternative sources, particularly shorter copular ones, are a very interesting concept which we shall return to when considering the account of sluicing proposed in this thesis (see Chapter 3), a more detailed analysis of the cross-linguistic debate on this phenomenon exceeds the space and time limitations of this thesis, with the reader instead being referred to the references above for a more detailed consideration.

(37) They wanted to hire someone who speaks a Balkan language, but I don't remember which Balkan language [~~that was~~].

(38) They wanted to hire someone who speaks a Balkan language, but I don't remember which Balkan language [~~he should speak~~].

Although the PF-Deletion Hypothesis - together with the Fit Condition - has been one of the most successful theories of sluicing to date and appears to account for a large amount of cross-linguistic variation within this phenomenon, its descriptive adequacy is, nevertheless, proving to have its limitations, as we shall see shortly, despite no other existing theories being able to better capture these data.

1.2.5 Pointer or Anaphor Account

We now move on to the last theory of sluicing we shall be discussing here, a sentence processing one. This is not a theory specifically designed based on the traits of sluicing per se, as the previous theories we saw had been, but was rather created for other types of ellipsis, such as VP-Ellipsis (Culicover & Jackendoff, 2005; Hardt, 1993; Martin & McElree, 2008; Tanenhaus & Carlson, 1990), and has been argued to also be applicable to sluicing (Martin & McElree, 2011). According to this pointer or anaphor account, the *e*-site does not contain structure of any kind at any point of the derivation. Instead, it is posited to contain a pointer, which guides the parser back to the antecedent to retrieve information, similarly to how some have argued for empty categories in the fields of *wh*-movement and pronoun resolution (Pickering & Barry, 1991; Pollard & Sag, 1994). Following the above discussion regarding the pros and cons of various theoretical approaches, we can immediately see the allure of such a structurally empty *e*-site hypothesis when dealing with instances of island repair under sluicing. On the other hand, it is less clear, at least at first glance, how this type of account would be able to secure connectivity effects.

One version of this approach, which we will also see in use when discussing P-stranding under sluicing in Polish, is along the lines of Culicover and Jackendoff (2005); Ginzburg and Sag (2000); Nykiel (2013); Sag and Nykiel (2011), a.o. This view considers memory mechanisms and their limitations to be the only governing factor in ellipsis resolution. That is to say, when a remnant is encountered, its semantic, syntactic and phonological features along with their relationship to those of the correlate are what make ellipsis resolution more or less acceptable. As such, it predicts memory-based effects to also affect ellipsis resolution, such as interference during correlate retrieval from similarly feature-coded phrases within the antecedent, which are not the target-correlate, however. Another similar prediction, which features prominently in the explanation of P-stranding in Polish, is that the discourse salience and feature informational complexity of the target and the remnant will affect ease of retrieval. Specifically, it is posited that the more complex the correlate, the easier the process of retrieval should be through a less complex remnant and vice versa (following Accessibility theory as proposed by Ariel (1990, 1994, 2001))[12]. Thus, when the correlate is a complex PP, e.g. *for which minister*, it is possible to refer back to it via a less complex

[12]It should be noted that it is not the net sum of complexity which they argue affects retrieval, as we might predict, e.g. complex correlate + complex remnant > complex correlate > / = complex remnant. Instead, following Ariel's Accessibility Theory, there appears to be a complexity trade-off between the referent and referee in anaphora resolution, with no more or less information being expressed than necessary; i.e. a more complex and, thereby, discourse prominent correlate is proposed to be more easily accessed through a less complex remnant, and vice versa, a less complex correlate will necessitate a more complex remnant in order to be activated. This appears to follow from the 'repeated name penalty' that some have argued for in sentence processing, i.e. that discourse-prominent referrent repetition during parsing (e.g. *John went to the market after John bought a cake* vs *John and Mary went to the market after John bought a cake*) may in fact hinder parsing (Almor, 1999; Garrod, Freudenthal, & Boyle, 1994; Gordon, Grosz, & Gilliom, 1993; Gordon, Hendrick, Ledoux, & Yang, 1999; Swaab, Camblin, & Gordon, 2004).

P-stranded *which*-NP, e.g. *which minister*. On the other hand, a simplex PP (*for someone*) is much less easily referred back to via a P-stranded *wh*-XP (*who*). This type of reasoning generates a rather straightforward cross-linguistic prediction with respect to acceptability gradients in sluicing and ellipsis more generally; following Ariel's Accessibility Theory, in all languages we should find that complex correlates encourage P-stranding sluicing remnants, whereas simplex correlates do not. We would, furthermore, anticipate that P-stranding a remnant should in fact be *more* acceptable than P-pied-piping one, provided the correlate is complex enough, with the P being superfluous. Another straightforward prediction would be that cross-linguistically we should find that informationally rich or complex correlates pattern with informationally poor or simplex remnants and vice versa. This last prediction, in particular, however presents a stumbling block for this theory. Following the above logic, the most successful correlate and remnant pairings should be a complex correlate (*some minister*) referred back to by a simplex remnant (*who*) and, conversely, a simplex correlate (*someone*) referred back to by a complex remnant (*which minister*), something which is not true. Indeed, these pairings have been argued to be impossible (see also M. V. Barros (2013); Dayal and Schwarzschild (2010); Jacobson (2016) for more in-depth discussions).

This last issue aside, however, this theory might be a little too restrictive, as we shall see in section 1.3.3. If one were to be more lenient about this information trade-off between correlate and remnant, one can see how this idea of more complex correlates encouraging P-stranding could also perhaps be related back to the way Hofmeister too argues for complexity playing a significant role in memory retrieval mechanisms and filler-gap dependency resolution (Hofmeister, 2011; Hofmeister & Sag, 2010). Specifically, more complex targets in terms of syntactic and semantic information are argued to be more easily and quickly retrieved in self-paced reading studies compared to less complex targets (Hofmeister, 2011). Additionally, there appears to be a gradient in how acceptably overt islands can be escaped, based on the complexity and size of the moved element; specifically, the larger the size and complexity of the moved element, the easier or more acceptable the movement out of an island (Hofmeister & Sag, 2010). In both these situations, complexity in terms of semantic and syntactic information translates to easier processing. When neither correlate nor remnant are complex, then anaphora resolution becomes harder. If we follow this reasoning, one could simply predict that more complex *wh*-remnants and more complex correlates are more easily P-strandable in general, i.e. following less the logic of Ariel (1990, 1994, 2001) and more that of Hofmeister (2011); Hofmeister and Sag (2010), and arguing for a combined complexity advantage, ignoring the repeated name penalty. All this having been said, however, whichever prediction we accept, due to a lack of more strictly controlled and adequately powered experiments, there is currently little evidence to support this type of hypothesis or generalisation in any language, contrary to what appears to be argued for in Nykiel (2013), as we shall see.

Another problem for this approach is that although it captures island amelioration under regular sluicing by default, it is unclear why some, but not all, islands are ameliorated, such as islands under sprouting ((21), repeated below).

(21) **Sprouting Islands**

Wh-Island; Subject Island; Complex NP Island

a. * Louie was thinking how to paint, but it was not clear what.

b. * That the guests would want to drink was likely, but we did not know what.

c. * They wanted to hire someone who drew well, but it was not clear what.

Finally, the way the theory proposes to capture connectivity effects appears rather brute-forced: in Sag and Nykiel (2011), in order to capture case-matching between correlate and remnant, a seemingly random post-hoc rule is enforced for sluicing whereby the remnant must always match the correlate in number, gender and case. When examining this rule, first of all it would appear as though its justification is non-existent other than to explain why the case-matching generalisation is predominantly observed cross-linguistically. In a similar fashion, one could simply argue that any number of rules could be added to the theory to capture any number of effects post-hoc, an inelegant solution to problems. Logical argument aside, in practice this general rule also fails to account for interesting mismatching effects which have been observed cross-linguistically for sluicing (and other forms of ellipsis), with respect to case- (see above, (24), and here, Chapter 2), number- and gender-marking, as well as voice. This feature of sluicing is something we will explore a little more in-depth in the next chapter and return to again throughout this thesis, with a few examples below of acceptable mismatches from English and Greek (39 - 42) (see also Merchant (2014) for similar mismatches under Greek nominal ellipsis).

(39) **Gender mismatching in English contrast sluicing**[13]

There always has to be a man and a woman on duty at airport security. Peter was there today, but I don't remember which woman.

(40) **Gender mismatching in Greek contrast sluicing**

O Petros sozi sinechos adespota zoa. Mechri
*Det.*M.NOM *Peter.*M.NOM *saves.*3SG *constantly* *stray.*N.ACC.PL *animals.*N.ACC.PL *until*

tora echi mazepsi dekapende skilous, alla den ksero poses
now *has.*3SG *collected* *fifteen* *dogs.*M.ACC.PL *but* NEG *know.*1SG *how.many.*F.ACC.PL

jates.
*cats.*F.ACC.PL

[13]Thank you to Klaus Abels for pointing these out.

'Peter is constantly saving stray animals. Up until now he has saved fifteen dogs, but I don't know how many cats.'

(41) **Number mismatching in English contrast sluicing**
Peter was on duty today at the airport, but I don't remember which other officers.

(42) **Number mismatching in Greek contrast sluicing**

I	filozoiki	proselave	ton	Petro,	alla	de
*Det.*F.NOM	*animal.rescue.center.*F.NOM	*hired.*3SG	*Det.*M.ACC	*Peter.*M.ACC	*but*	NEG

thimame	posous	allous	ethelondes.
*remember.*1SG	*how.many.*M.ACC.PL	*other.*M.ACC.PL	*volunteers.*M.ACC.PL

'The animal rescue center hired Peter, but I don't remember how many other volunteers.'

Even if we do not accept the above examples, however, following the strict case-matching rule between correlate and remnant, it is unclear why sprouting, lacking an overt correlate, need also bear a specific case, i.e. the case it would normally be assigned if the remnant's continuation were an overtly expressed copy of its antecedent clause (minus the *wh*-phrase).

Another version of this approach consists of a slightly different interpretation of the pointer mechanism from within the field of sentence processing and which works on a different level of cognitive analysis to the above (Marr, 1982). This is the approach of Martin and McElree (2011). For them, a pointer mechanism serves to explain how the parser must re-access the antecedent in memory in order to successfully resolve the sluice in real time. Their work predominantly deals with what exactly is being accessed, what restrictions govern this process, and what features may ease it, examining both sluicing and VP-ellipsis. With respect to the level of syntactic analysis of the *e*-site itself, in terms of structure, however, they appear to accept Merchant's PF-Deletion Hypothesis, citing that the remnant must indeed be created as part of a full CP whose IP is deleted through Merchant's *e*-GIVENness. As such, connectivity and anti-connectivity effects are taken to follow as in Merchant (2001, 2008). This is not unlike the approach I intend to take here, by combining PF-Deletion with a sentence processing approach, specifically that of considering sluicing to behave as a noisy channel. The point of divergence between Martin and McElree (2011)'s pointer account and the one argued for here is the degree of flexibility that they afford to the sluice as a whole. Specifically, in its current form, the former would not allow for a (gradient) acceptability of P-stranding or case-mismatching of remnant and correlate, as the search mechanism Martin and McElree (2011) argue for is a direct-access content-addressable one based solely on the cues provided by the sluice. These cues have been argued to be grammatical and are used to constrain the search set, disregarding all non-matching material. That is to say, only a sluice which grammatically matches the antecedent and exists in a relationship of mutual entailment with it will contain the cues necessary to successfully access the antecedent from memory, i.e. for a full PP antecedent, this would require a case-matching and P-pied-piping remnant. A P-stranded remnant or a case- or gender-mismatched remnant, on the other hand, would not provide correct, matching cues to the antecedent, reducing the antecedent search operation to a search set of 0, effectively rendering the structure unacceptable. In order to overcome this issue and capture differences in remnant acceptability not only in the same language, but also across different languages, this hybrid pointer approach would need to generate additional hypotheses on specific features which aid or inhibit antecedent recovery cross-linguistically beyond precisely matching cues. To the best of our knowledge, as the theory currently stands, these types of additional hypotheses have not yet been proposed. The theory proposed in this thesis, on the other hand, is not limited by these restrictions, as shall be made clear in due course.

Another important point to mention regarding evidence for or against this type of account, as with arguments for or against empty categories in *wh*-movement, is how often and easily studies mistakenly treat experimental evidence showing antecedent access at the *e*-site or gap site as evidence for or against phonologically null structure being created at that site (see e.g. Nykiel (2013)). Specifically, studies using ERPs, cross-modal priming and visual world-paradigms have all shown evidence that lexical, semantic and phonological levels of the antecedent are clearly re-accessed at the *e*-site in VP-ellipsis (Kaan, Wijnen, and Swaab (2004); Shapiro, Hestvik, Lesan, and Garcia (2003); Snider and Runner (2010), though cf. Kaan, Overfelt, Tromp, and Wijnen (2013)). Similarly, a self-paced reading study involving sluicing (Yoshida et al., 2013) showed evidence for rapid re-accessing of detailed, hierarchical binding relations within the antecedent at the sluice's *e*-site (although, see here Chapter 4 for evidence against the results of this study in particular). Although we agree that, undoubtedly, these results provide clear evidence for various linguistic levels of the antecedent being accessed at these points, we disagree that they are clear evidence either for a pointer mechanism or for structure building at the *e*-site, maintaining instead that it is practically impossible to distinguish the timing or other processing effects of one phenomenon from the other. As Phillips and Parker (2014) also persuasively point out, if these sorts of experiments have been considered inconclusive for similar arguments in the *wh*-movement and pronoun resolution literature with respect to phonologically null content at the *e*-site (Kempen, 2010; Phillips & Wagers, 2007), then it is unclear why they should be considered conclusive here.

1.3 P-Stranding under Sluicing in Non-P-Stranding Languages

1.3.1 Two Classes of Explanation

Having covered the major theoretical and sentence processing approaches to sluicing, we can now consider in more detail those intriguing cases where it has been argued that in certain overtly non-P-stranding languages a P-less sluice remnant is nevertheless allowed with a PP correlate. In other words we appear to have 'P-stranding' under sluicing in these languages despite its being disallowed overtly[14]. The cross-linguistic scene, however, as we shall see, is by no means homogeneous or consistent, with respect to either the data collection processes used, the types of explanation proposed, or the degree of cross-linguistic generalisability and general validity of these explanations.

Overall, all such documented exceptions to the PSG can be broadly separated into two categories based on the type of explanation proposed by the authors. In the first category, we find those studies whose authors continue to favour a PF-Deletion approach, suggesting either tweaking the theory slightly to accommodate for these exceptions based on linguistic idiosyncrasies or that these seemingly defiant remnants do not in fact constitute true counter-examples to the PSG as it stands. In the second, much smaller, category, we find those studies considering P-stranding instances to be true counter-arguments not only for a PF-Deletion Hypothesis, but any theory, in general, positing structure of any level at the *e*-site. Instead, these studies propose base generation of the remnant coupled with a purely sentence processing approach, such as the pointer account mentioned previously, in order to provide a rationale for island amelioration together with such P-stranding exceptions. To capture connectivity effects, additional identity requirements between antecedent and remnant are then independently imposed. Whichever category we take, however, it rapidly becomes clear that these explanations account for little cross-linguistic data beyond the language in question.

Before we move on to review examples of each of these categories, however, it should first be made clear that for the vast majority of these cases, the judgements given are purely anecdotal, i.e. they are either the judgements of the authors themselves, of one or two fellow linguist colleagues, or perhaps simply one naïve speaker, without any experimental testing. They are, furthermore, for the largest part, not graded in acceptability, instead

[14]One could arguably debate here whether this phenomenon should indeed be addressed as 'P-stranding' under sluicing, given we have no overt verb or stranded P, with this terminology thus showing a theoretical inclination towards a movement-based hypothesis of sluicing, or whether it should instead be termed simply 'P-omission' or 'P-deletion'. Although I agree with this being an issue and will return to it later on, for reasons of consistency with the existing literature discussed below, as well as the theoretical approaches discussed above, for the time being I will continue to refer to this phenomenon as 'P-stranding' under sluicing, without necessarily implying a movement-and-deletion hypothesis unless otherwise stated.

treating judgements as acceptable, unacceptable and perhaps some grey area denoted by '?'/'??'. I consider these cases interesting and report them here as important indications of there perhaps being something intriguing occurring in these languages. Given this lack of adequate experimental evidence, however, I do not consider them as clear evidence for or against P-stranding actually being acceptable in any of these languages, nor, more importantly perhaps, as clear indications of noteworthy language-specific sluicing behaviours tied to linguistic idiosyncrasies, as some have argued for. Each of these cases is clearly important and deserves further investigation, with precise, controlled studies based on theoretically-motivated and statistically disconfirmable hypotheses. For the few cases that actually involve experimental testing, I will briefly evaluate their experimental methods, hypotheses and conclusions and consider them when analysing my own data in an effort to create a more realistic, rounded and cross-linguistically valid explanation and set of predictions.

1.3.2 PF-Deletion in spite of P-Stranding

1.3.2.1 P-Stranding thanks to Alternative Sources

Beginning with the first category of explanation, there is a group of studies involving non-case-marking languages which argue that these apparent exceptions to the PSG are not true instances of P-stranding. Instead, what appears to be a P-stranded remnant is argued to hail from an alternative source which does not contain P-stranding, as such not constituting a true counter-example to PF-Deletion as a theory. For these cases, there are two main alternative sources proposed.

1.3.2.1.1 Short Copular Source

The first alternative source proposed are shorter or longer copular structures[15] along the lines of Erteschik-Shir (1977), Merchant (1998) and Van Craenenbroeck (2004, 2010b), such as (43a) having the copular structure in (43b) as its source instead of the P-stranding one in (43c). These alternatively sourced instances of ellipsis have been termed cases of 'pseudosluicing' by Merchant (1998), i.e. not 'real' instances of sluicing, given this alternative source.

(43) a. Maria danced with someone, but I don't remember who.
b. Maria danced with someone, but I don't remember who that/it was.
c. Maria danced with someone, but I don't remember who she danced with.

[15]These copulas have frequently been termed 'cleft' sources in the literature, even though they are not always clear clefts.

The most well-known argument for a cleft-like source involves the languages of Brazilian Portuguese and Spanish. Brazilian Portuguese (henceforth 'BP'), a non-P-stranding language (44), was first documented to allow P-stranding under sluicing by Almeida and Yoshida (2007), as in (45).

(44) *(Com) quem$_i$ que a Maria dançou t$_i$?
with who$_i$ that the Maria danced t$_i$
'With whom did Maria dance?'

(45) A Maria dançou com alguém, mas eu não lembro (com) quem$_i$
the Maria danced with someone but I not remember with who$_i$
~~a Maria dançou t$_i$~~.
~~the Maria danced t$_i$~~
'Maria danced with someone, but I don't remember (with) who.'

In most other respects, ellipsis in BP would appear to behave as expected; sluicing seems able to ameliorate islands, as opposed to VP-Ellipsis, with the latter also not appearing to allow P-stranding. As such, it could not simply be argued that - for some reason - PPs are PF-islands in BP, nor that the amelioration of all islands is an inherent property of all forms of ellipsis in BP.

The possibility that the sluicing example in (45) may actually have a cleft source, i.e. that it is in fact a pseudosluice (Merchant, 1998), was originally examined and rejected by Almeida and Yoshida via a series of diagnostics which Merchant (2001) himself uses to show that sluicing in English cannot always be reduced to pseudosluicing. Although I will mention these in brief, the logic behind using these diagnostics here is not always entirely sound (see also Van Craenenbroeck (2010a)). The first of these is based on prosody. Specifically, it is argued that in BP, as in English, different contours are used for clefts and sluices, with the former having an emphasis on the copula, and the latter on the *wh*-remnant. The second diagnostic involves the pattern of 'mention some' and 'else' modifications. Specifically, cleft pivots are compatible only with a 'mention-all' interpretation and thus cannot accept modifiers requiring a 'mention-some' interpretation, e.g. *for instance* or *such as*. These are perfectly acceptably allowed to modify a sluice remnant, however. For similar reasons, although the modifier 'else' can be applied to a sluice remnant, apparently it cannot be to cleft pivots. [16]

[16] An additional diagnostic which was also used concerns the behaviour of aggressively non-D-linked *wh*-phrases; although these are argued to be acceptable in clefts, they are apparently impossible to occur as a sluice remnant (1):

(1) A Maria dançou com um outro cara ontem à noite ...
the Maria danced with one other guy yesterday at night

At this point, it should be mentioned that the validity of these diagnostics to prove the specific point is questionable. What these comparisons show is that, as in English, BP sluicing is simply not *always* reducible to pseudosluicing, not that pseudosluicing is impossible as a sluice source (see Van Craenenbroeck (2010a) for an extensive analysis of this). As such, this does not mean to say that in instances of sluicing with P-stranding these cannot have an alternative, copular source. For this reason, it is only relevant to show here that a cleft source is impossible for P-stranded remnants. The most convincing argument against a cleft-source analysis for BP, hence, is that, according to Almeida and Yoshida (2007), clefts do not actually appear to allow P-stranding (46d)[17].

(46) A Maria dançou com alguém, mas . . .
the Maria danced with someone but
'Maria danced with someone, but . . .'

a. eu não sei com quem.
I not know with who
'I don't know with who.'

b. eu não sei com quem foi.
I not know with who was

'Maria danced with another guy last night . . . '

a. eu só queria saber (com) quem!
I only wanted know with who
'I wish I knew (with) who!'

b. eu só queria saber *(com) quem diabos foi!
I only wanted know with who devils was
'I wish I knew (with) who the hell it was!'

c. * eu só queria saber (com) quem diabos!
I only wanted know with who devils
'I wish I knew (with) who the hell!'

That having been said, I do not consider this an accurate argument as this diagnostic in its original Merchant (2001) iteration has been shown to be a non-sequitur (Van Craenenbroeck, 2010a). It is not enough here to show that sluices cannot occur with aggressively non-D-linked *wh*-phrases but that clefts can; it must also be shown that these aggressively non-D-linked *wh*-phrases are also impossible with *wh*-questions, something which is not necessarily true. Using an English example (2), if sluices not being able to occur with aggressively non-D-linked phrases is a valid argument against their having clefts as a source, then it is also an argument against their having *wh*-questions as a source, since the two are not distinguishable in this regard.

(2) a. I don't know who the hell it was that Maria kissed.
b. I don't know who the hell Maria kissed.
c. * I don't know who the hell.

[17]Although it should be noted that (46d) is judged as marginal and not completely unacceptable, as opposed to overt P-stranding (44).

'I don't know with who (it) was.'

c. eu não sei quem.
I not know who
'I don't know who.'

d. ?? eu não sei quem foi.
I not know who was
'I don't know who (it) was.'

As a result, the authors exclude the possibility that P-stranding in BP has a cleft-source. Instead, their explanation of this phenomenon is that in BP and similar, but not all, languages, P-stranding should be considered a PF-violation, following, for instance, Aoun, Hornstein, Lightfoot, and Weinberg (1987), resulting in P-stranding being ameliorated under sluicing within a PF-Deletion framework. For all the other overtly non-P-stranding languages which do not show this amelioration, it is argued that the constraint on moving out of the PP must be a constraint on *wh*-movement, for instance with PP being a phase, and hence deletion at PF not affecting it (Abels, 2003). As with many such approaches, however, by treating non-P-stranding languages which exhibit P-stranding under sluicing as somehow special without clear explanation as to why, it is unclear why PPs would behave as islands at PF for certain non-P-stranding languages only and as a restriction on *wh*-movement for others; it is also not immediately clear why PF-islands should be ameliorated under sluicing but not any other forms of ellipsis targeting constituents larger than PP, nor what theoretically makes islands be distinguished into PF- and non-PF-islands.

Almeida and Yoshida's reasoning is later questioned by a related study by Rodrigues, Nevins, and Vicente (2009) who examine both BP and Spanish, another apparently PSG-defying language. As opposed to Almeida and Yoshida (2007), however, they do not consider these to be true instances of P-stranding under sluicing, and as such do not constitute real counter-examples to a PF-Deletion approach. Instead, it is proposed that Romance languages exhibiting this behaviour have two potential sluicing sources; one involving a regular IP which includes *wh*-movement; and one involving a longer copular source with no such movement. Both of these sources are proposed to comply with the predictions of the PF-Deletion framework, with P-stranding under sluicing being allowed only in those cases where a copular source is also available.

At first glance, this explanation may appear inconsistent with the claim made by Almeida and Yoshida that BP does not allow P-stranding under clefting (46d). However, Rodrigues et al. posit a different, more elaborate source, one including a specificational copula, i.e. followed by a DP which in turn is modified by a restrictive relative clause (as in (47)). It is within this RC that the required *wh*-movement occurs with pied-piping of the P, followed by subsequent deletion of the copula at PF, leaving a bare cleft pivot at PF and giving the illusion of P-stranding.

(46d) Almeida and Yoshida cleft source (repeated):

A Maria dançou com alguém, mas eu não sei ??(com) quem foi.
the Maria danced with someone but I not know with who was

'Maria danced with someone, but I don't know (with) who (it) was.'

(47) Rodrigues et al. cleft source:

A Maria dançou com alguém, mas eu não sei quem é com a qual a Maria dançou.
the Maria danced with someone but I not know who is with the that the Maria danced

'Maria danced with someone, but I don't know who it is with which she danced.'

Although in the interest of brevity, not all of these will be covered here, a number of other important facts are also given in support of this proposal, such as that despite multiple sluicing being allowed in both BP and Spanish, neither shows P-stranding under multiple sluicing (48). Given Almeida and Yoshida (2007) propose that extraction out of PPs at PF is generally allowed in languages showing exceptional P-stranding behaviour, this would be an incongruous observation for their approach. The explanation proposed for this observation also consitutes in itself an argument for a cleft analysis; specifically, it is based on Lasnik (2006; reiterated in 2014) who analyses multiple sluicing in English as single *wh*-fronting accompanied by rightward movement of the second *wh*-phrase plus deletion of the IP (49). Following the same logic, Rodrigues et al. (2009) propose that BP and Spanish do not allow P-stranding for the second *wh*-phrase as it is considered to extrapose rightwards, a movement which inherently does not allow P-stranding.

(48) a. Spanish

Ella habló con alguien sobre algo, pero no sé *(con) quién *(sobre) qué.
She talked with someone about something but not know with who about what

'She talked with someone about something, but I do not know (about) what (with) who.'

b. Brazilian Portuguese

Ela falou sobre alguma coisa para alguém, mas eu não sei *(sobre) o que
She talked about some thing to someone but I not know about the what
*(para) quem.
to who

'She talked to someone about something, but I do not know (about) what (to) who.'

(49) a. Maria argued with someone about something, but I don't remember (with) who *(about) what.

b. ... but I don't remember [$_{CP}$who [$_{IP}$~~Maria argued [with~~ t_1]]] [t_2]] [about what]].

Regarding the first *wh*-XP, if this were to appear preposition-less, two things would presumably occur, a) there would be a specificational copula source, b) the second *wh*-phrase would be contained within the RC of this source and would have to move rightwards out of it in order to escape deletion, thus crossing a further clausal boundary and violating the Right Roof Constraint (Ross, 1967)[18]. Hence, neither *wh*-phrase is allowed to appear bare[19].

Further instances where P-Stranding patterns with clefts in terms of acceptability include 'else' modification (50), since in BP this is allowed for both clefts (as defined by Rodrigues et al. (2009), i.e. with a specificational copula; cf. Almeida and Yoshida's argument against 'else' modification for clefts) and their supposed P-Stranding sluice derivatives. Spanish, on the other hand, as is the case in English, does not allow such modification on either overt clefts or P-Stranded sluices.

(50) a. O João saiu com a Maria mas eu não sei quem mais.
The João went.out with the Maria but I NEG *know.*1SG *who else.*
'João went out with Maria, but I don't know who else.'

b. Me fala quem mais é que você quer convidar para sua festa.
Me tell who else is that you want invite to your party.
'Tell me who else it is that you want to invite to your party.'

[18]The Right Roof Constraint states that an element cannot move rightward out of the clause in which it originates. In this case, the constraint would have to apply prior to PF and not be possible to ameliorate under sluicing, something which, however, is not clearly testable.

[19]It should be noted, however, that it is unclear whether this is the correct interpretation of multiple sluicing; cf., for instance, Abels and Dayal (2017); Merchant (2001); Richards (1997, 2001).

Based on these, amongst other, arguments[20] Rodrigues et al. (2009) conclude that BP and Spanish P-Stranded remnants are only apparently contradictory to the PSG. Instead, an amended version of said generalisation is proposed ((51), their (69)):

(51) *Form Identity Generalization II: P-stranding (revised)*
For any syntactic configuration C, if P-stranding is banned in C in non-elliptical environments, it will also be banned in C under sluicing.

What this essentially implies is that a sluice source which involves P-stranding should not result in P-less remnants being acceptable with PP-correlates. Instead, if there is some conceivable sluice source that allows P-stranding, then it should be possible to have a P-less sluice remnant with a PP-correlate even when overt *wh*-movement does not allow P-stranding in this language.

1.3.2.1.2 Resumption

The other alternative source proposed for P-stranding under sluicing involves a resumptive pronoun (Wang (2007), AlShaalan and Abels (2019), e.g. (52) from Saudi Arabic).

(52) al-qana alʾawlā sawt maqābla maʿ wazīr saʿūdī bas nasīt ʾay
*the-channel the-first made.*3FSG. *interview with minister saudi but forget.*1SG *which*
wazīr ~~sawt maqābla ma-ʿh~~
*minister made.*3FSG. *interview with-him*
'Channel 1 did an interview with a Saudi minister, but I forgot which minister [~~Channel 1 did an interview with him~~].'

The most convincing argument for this source comes from AlShaalan and Abels (2019) for Saudi Arabic (SA). Based on Molimpakis (2016a, 2016b) and following the same experimental process outlined here in Chapter 2, AlShaalan and Abels (2019) ran a similar series of experiments modified for SA. These involved large-scale web-based acceptability judgements using a Likert scale rating system from 1 (least acceptable) to 7 (most acceptable). It should also be noted that this is the only other language to have been examined with

[20] It should be mentioned that Rodrigues et al. (2009) also contributed an additional argument against using aggressively non-D-linked *wh*-phrases (or *RPIs* as they are termed by Rodrigues et al.) as a valid argument against P-stranding having a copular source, aside from the fact that is is a non-sequitur to begin with. Specifically, they argue that the disallowance of RPIs with (apparently) stranded remnants, even though they are allowed with clefts, may in fact be due to an altogether different, phonological reason related to sentence accenting. Specifically, RPIs, they argue, cannot be the last pronounced sentential element (as in English) since a) they are not a category which can receive accenting (following Gussenhoven's Sentence Accent Assignment Rule, Gussenhoven (1984)); b) this means they must belong to another focus domain which receives accent, namely the following material; c) there is no following overtly expressed material (Sprouse, 2006).

adequately powered and controlled experiments aside from our investigations here of Greek and German.

To explain these experiments, we must take into consideration that the creation of *wh*-questions in SA can be done in three distinct ways: via regular *wh*-movement (53a), through a cleft design (53b) or with a resumptive pronoun (53c), (their (2)).

(53) a. ʾay bant šaft-ī __ ?
*which girl see.*2FS.
'Which girl did you see?'

b. ʾay bant hay aly šaft-ī-hā?
*which girl she that see.*2FS.*her*
'Which girl is it that you saw?'

c. ʾay bant šaft-ī-hā?
*which girl see.*2FS.*-her*
'Which girl did you see her?'

The experiments in this study were designed around the fact that each of these three methods behaves differently with respect to the factors of *wh*-type, contrastive focus and P-pied-piping. Importantly, out of the three, only regular *wh*-movement requires P-pied-piping, with it being optional for both cleft-structures and resumption structures. Based on these different behaviours, a series of acceptability judgment studies were conducted to disentangle whether P-stranding allowance under sluicing in SA can possibly be linked to a cleft source or a resumptive pronoun source.

In the first of their experiments, a cleft source was blocked by utilising contrast sluicing, as copulas and 'else' modifications have been argued to be disallowed cross-linguistically with contrast sluicing, thus leaving resumption and regular *wh*-movement as potential sources[21]. The results showed that despite this cleft source being blocked, P-stranding was still acceptable under sluicing (Mean: 5.06; SD: 1.91), similarly to P-pied-piping (Mean: 5.38; SD: 1.88), and much more so than with overt *wh*-movement (P-stranding Mean: 2.26; SD: 1.72; P-pied-piping Mean: 5.21; SD: 2.04). This would indicate that for these cases of sluicing, either resumption is being used as an alternative, acceptable source, or these are in fact true cases of P-stranding with *wh*-movement as their source. In another experiment, overt P-stranding with resumption was compared to P-stranding under sluicing, finding the two to be similarly highly acceptable (Mean overt: 5.80; Mean sluicing: 5.78), with no significant difference between them. There was also no difference between P-stranding and P-pied-piping, either overtly (P-stranding Mean: 5.80; SD: 1.49; P-pied-piping Mean: 5.47;

[21]This unacceptability was also confirmed experimentally by including such items as fillers in their last experiment.

SD: 1.77) or under sluicing (P-stranding Mean: 5.78; SD: 1.47; P-pied-piping Mean: 5.97; SD: 1.40), indicating that a) resumption was just as acceptable with P-stranding as it was with P-pied-piping; and b) P-stranding and P-pied-piping were equally acceptable under sluicing. Finally, in the last experiment, they created a condition where both cleft and resumptive pronoun sources were blocked, leaving only *wh*-movement as a possible source for P-stranding under sluicing. This was done by utilising the adverb *when*, which cannot pattern with either resumption or a copula, and comparing it to *where*, which can pattern with either. Their results showed a significant difference between the two adverbs, with P-stranding with *where* being more acceptable than P-stranding with *when* (Mean for *where*: 6.15; Mean for *when*: 4.62). Interestingly, however, these results also indicated that even with the adverb *when*, i.e. when both these alternative sources were being blocked, P-stranding was still more acceptable under sluicing than it was overtly (Mean: 2.53). That having been said, it was not as acceptable as either P-pied-piping overtly or under sluicing (Mean: 5.81), nor was it as acceptable as P-stranding was when an alternative, resumptive pronoun source was available (i.e. for *where*). Based on these findings, AlShaalan and Abels conclude that P-stranding under sluicing in SA does not constitute a true counter-example to a PF-Deletion theory, but that it has resumption as an alternative, acceptable source at its core. It should be noted though that, if this is the full story, it is not made clear how or why P-stranding is still much more acceptable under sluicing in SA than overtly when all alternative sources are blocked, even if it is not as acceptable as when they are not; acknowledging this, the authors leave this question open for further investigation. I believe this last finding, in particular, shall prove very interesting when I propose an alternative explanation for P-stranding under sluicing in such situations in the coming chapters.

1.3.2.2 P-Stranding without Alternative Source

For languages that can, quite elegantly, entertain an alternative source for apparent P-stranding under sluicing, it appears as though the revised PF-Deletion Hypothesis and its treatment of PPs under sluicing need not be altered in any significant way. However, as we shall see, not all languages are capable of such alternative sources. In this situation, the cross-linguistic landscape appears much more muddled and incohesive with respect to the various alternative explanations proposed.

Still working within a PF-Deletion framework, authors who have - anecdotally - reported evidence of P-stranding under sluicing in such, overtly non-P-stranding languages mostly propose explanations tied to the idiosyncrasies of the language in question. These approaches are also, quite often, less than convincing in their argumentation given that some of these idiosyncrasies are shared with other languages, yet this sluicing behaviour does not appear to generalise to them. One such approach is Sato's (2011) explanation of P-stranding in Bahasa Indonesian (BI). After excluding the possibility of pseudosluicing as an alternative source for

P-stranding in this language, Sato proposes that P-Stranding is acceptable under sluicing due to a PF-interface repair process based on the concept of [+wh] feature percolation (Chomsky, 1972; Lasnik, 2005). Specifically, depending on the language, a [+wh] DP may copy this interrogative feature up to its governing PP. This PP thus becomes [+wh]-marked and thereby the closest interrogative element to C, resulting in the full PP being attracted to C under regular *wh*-movement, our familiar P-pied-piping. The DP itself, on the other hand, cannot move out of the PP in this case due to its violating the A-over-A type constraint (Chomsky, 1966). This percolation process is optional or mandatory depending on the language, its verification, however, is posited by Sato (2011) to be a PF-process. When the [+wh] feature percolation onto PP is obligatory, as for instance in BI and French, then splitting this PP under regular *wh*-movement would indicate a 'failure of the [+wh] feature to percolate at the PP level' (Sato, 2011, p. 366), as otherwise the [+wh] DP would be violating the A-over-A type constraint. However, deletion of the offending PP at PF via IP-ellipsis would mean the PF-rule would have nothing to apply to, resulting in the structure's amelioration. This idea appears very much in line with that of ungrammatical feature deletion being the driving force behind island amelioration under sluicing (Ross, 1969)[22].

Of course, the next logical question would be why this is not also true for all languages enforcing P-pied-piping under regular *wh*-movement. To account for this distinction, Sato calls upon other, independent syntactic violations occurring earlier in the derivation which, he argues, cannot be salvaged by PF-Deletion given they lead to the derivation crashing before reaching PF. Without going into too much syntactic detail here, the phenomenon of Determiner-Preposition coalescence is proposed to be of pivotal importance. This attribute, distinguishing BI from French for instance, is a characteristic of certain Romance languages where D and P can merge into a suppletive form (Law, 1998; Van Riemsdijk, 1998), e.g. French: ***du***/de le; ***aux***/à les; etc. German: ***am***/an dem; ***aufs***/auf das; ***zur***/zu der; etc. Sato argues that DP coalescence cannot co-occur with P-stranding, hence all DP-coalescent languages should be unable to allow P-Stranding under sluicing. However, from Spanish (Rodrigues et al., 2009) and French (Rodrigues et al. (2009); Ott and Therrien (2018) for Ontario French), to Italian (Abels, 2003), to Greek and German (Chapter 2), there is ample evidence that this does not appear to be the case, with these DP-coalescent languages indeed appearing to allow P-stranding under sluicing. An alternative explanation should, therefore, be sought elsewhere.

This concept of PPs being salvageable within a PF-Deletion framework as some form of

[22]It is interesting to point out, however, that apparently even within languages there is some parameterisation between different prepositions. Sato proposes, for instance, that certain Ps which do not appear to be strandable even in English - a normally P-Stranding language - behave thus because they have the *lexical* property of obligatory [+wh] percolation at PP. Under Sato's current proposal, these should be acceptably strandable under sluicing, a prediction which indeed appears to be borne out (Chung et al. (1995); Rosen (1976)). On the other hand, it is unclear why this should be the case for certain Ps only, and Sato himself calls this a 'random' lexical property.

PF-island recurs in Leung (2014) for Emirati Arabic. Specifically, Leung (2014) argues that if a language exhibits P-stranding under sluicing, then PPs should be considered PF-islands in this language. This somewhat cyclic argument, however, fails to explain why PPs are PF-islands in the first place for some, but not other, overtly non-P-stranding languages.

Another approach is taken by Stjepanović (2008) for Serbo-Croatian (SC). Excluding the possibility of all alternative sources, due in part to remnant case-marking[23], Stjepanović (2008) argues that P-stranding in SC must be a result of regular *wh*-movement followed by deletion. The matter, however, is convoluted by the apparent behaviour of multiple coordinated PPs under sluicing. Specifically, in (54), *koje* and *kojeg*, two *wh*-remnants of differing morphological case, may appear P-less and coordinated, whilst their correlates are also two coordinated PPs.

(54) Petar je sakrio igračku ispod jedne stolice i pored jednog zida, ali ne
Petar is hidden toy under one chair.GEN. and beside one wall.GEN but not
znam (ispod) koje stolice i (pored) kojeg zida.
*know.*1SG *under which chair.GEN and beside which wall.GEN.*
'Petar hid the toy under a chair and beside a wall, but I don't know which chair and which wall.'

If this is to be considered a true case of sluicing, then one would expect that the P-less remnants originated within PPs and subsequently lost their P at some point during the derivation. Stjepanović argues that two options present themselves at this point, the interpretational result of each of which is crucial to understanding the true nature of the elided structure in (54). The first option is to consider that these are coordinated PPs behaving as a single constituent, simultaneously moving and losing their Ps at some point during the derivation. Problematically, however, no existing theory of movement can actually sustain such a coordinated operation whilst dropping both the PPs' Ps. From an interpretational point of view, if this is true, then one would expect them to behave as a single entity and be able to be referred to as such, i.e. the remnants in (54) would be able to be interpreted as a single place. The second option is that these are not coordinated PPs, but coordinated CPs which behave independently of one another, with the PP embedded within each losing its P somewhere along the derivation, *independently* of the other. This option is supported by current theories of movement and sluicing, however runs up against the obvious problem of P-stranding not being allowed in this language. From a semantic point of view, if this is the case, then there should not be a reading available for (54) as a single entity. In other words, 'under a chair and beside a wall' would have to always be interpreted as two separate places, one being under chair and the other beside a wall.

[23]Specifically, cleft clause pivots are always marked by Nominative in SC, whereas the P-stranded sluice remnant appears in the case it would normally be assigned by the P, a *non*-Nominative case.

Adjudicating between the two, Stjepanović (2008) states that, semantically, it is in fact possible for the sluice in (54) to be interpreted as a single place, consistent with the first derivational option. As such, it is thought that although P-stranding is apparently available under sluicing in SC, this is achieved in a way which is not supported by current theories of movement, since such theories would exclude a single entity reading. This leads her to the conclusion that these are not in fact true cases of P-stranding, at least not in the usual sense of P-stranding as a result of regular *wh*-movement. As a result, they do not present an issue for the PF-Deletion Hypothesis. Although an explanation is not provided for the real source of these effects, it is putatively thought to be a 'post-syntactic phenomenon, occurring possibly at PF', as any syntactic account has been 'exhaust[ed]' (Stjepanović, 2008, p. 188).

These are simply some of many such alternative approaches, none of which, however, has involved large-scale, controlled acceptability ratings nor proposes a cohesive, cross-linguistically generalisable explanation which can account for all the data we have to hand. For a more comprehensive list of languages and explanations, please see table 1.2 at the end of this chapter.

1.3.3 P-Stranding therefore No PF-Deletion

Moving away from PF-Deletion and a syntactic account of sluicing, in general, towards an approach based purely on sentence processing, we arrive at Nykiel's (2013; 2014; 2015) and Sag and Nykiel's (2011) treatises of Polish sluicing. As with all other languages in this section, Polish has been proposed to be an apparently P-stranding language under sluicing, despite not being one overtly. Originally, a cleft structure had been proposed as the acceptable alternative source of apparent P-stranding in Polish by Szczegielniak (2006), however this was argued to, in fact, be impossible by Nykiel (2013). Specifically, it appears as though clefts in Polish do not allow P-stranding overtly, for either complex or simplex *wh*-phrases, hence such an explanation would be pointless.

Instead, Nykiel (2013) argues for a pointer mechanism approach to explain Polish data, eschewing all structure within the *e*-site. It should be made clear at this point, as was also done in section 1.2.5, that a pointer account does not inherently rule out structure at the *e*-site, but that this depends on the researcher's approach. E.g. in Martin and McElree (2008), it is argued that a pointer mechanism applies at the level of sentence processing, but that a PF-Deletion approach should be held to explain connectivity and anti-connectivity effects. Nykiel (2013); Sag and Nykiel (2011), on the other hand, argue for *no* syntactically rich structure at the *e*-site at any point, instead positing a silent pointer therein, which serves to direct the parser back to the antecedent in order to gather all the semantic information it requires to accurately retrieve the sentence's meaning. Sluicing, as such, should behave as a type of anaphora, affected by the same working memory mechanisms, complete with their limitations. To remind the reader, in section 1.2.5 we examined two such mechanisms and their predictions for such working-memory-driven accounts. The first was Ariel's Accessibility Theory which involves an information trade-off between antecedent and referent. Based on this, it is predicted that the more complex the antecedent is, the less complex the referent should be and vice versa. The second mechanism follows Hofmeister (2011) in predicting that it is the total complexity of antecedent and referent together that affect ease of parsing, with greater overall complexity leading to easier parsing.

Out of these two, Nykiel (2013) argues for the first mechanism to underlie P-stranding under sluicing in Polish. This is supported, the author argues, by a preference for *which*-NPs as P-stranded remnants in Polish over *wh*-XPs, based on a series of Polish sluicing studies presented in Nykiel (2013). Although the validity of these experiments is questionable at best, something I will return to shortly, let us assume for the moment that *which*-NPs are indeed preferable over *wh*-XPs in allowing P-stranding. The explanation for this is that *which*-NPs typically make reference to a lexically rich NP correlate (e.g. *some minister*), whereas *wh*-XPs would typically refer back to a lexically poor indefinite (e.g. *someone*)[24]. Following Ariel's Accessibility Theory, due to *which*-NP remnants referring back to a complex correlate, they can also acceptably appear as less complex, thereby becoming better-suited to referring back to complex correlates. One way of appearing less complex, for instance, would be to appear as remnants without their NP. P-stranding is viewed through the same lens here and considered to make *which*-NP remnants less complex and thereby better-suited to refer back to complex PP correlates. According to this information trade-off logic, we could make two predictions. Firstly, by being less complex, P-stranded remnants should be better than P-pied-piped remnants when referring to complex PP correlates; secondly, *which*-NP remnants appearing without a P or an NP should be the most acceptable remnants for complex correlates. This is not, however, what Nykiel (2013) found. Her experiment results showed in order of acceptability preference that P-pied-piping *which*-NPs were the most acceptable, followed by P-pied-piping *wh*-XPs, in turn followed by P-stranding *which*-NPs and then P-stranding *wh*-XPs. Confusingly, this ordering of P-pied-piped remnants over P-stranded remnants is apparently explained via the theory of Nairne (2006), whereby greater feature overlap between a correlate and its retrieval cue lead to better accessibility; this translates to more complex correlates (*for some minister*) being more readily referred back to via more complex remnants (*which minister*), as they share multiple phonetic, semantic and syntactic features and this, in turn, makes P-stranding more acceptable than when dealing with simplex correlates and remnants. Given that the most feature overlap would result from a P-pied-piping remnant (*for which minister*) and complex correlate (*for some minister*) this appears completely contrary to Ariel's Accessibility Theory and the logic

[24]It is unclear here why this pairing is considered to be in accordance with Ariel's Accessibility Theory; the Theory would technically predict the opposite pattern, with complex correlates being referred back to by simplex remnants and vice versa, as also predicted in section 1.2.5, however this is not touched upon.

previously laid out by Nykiel. An explanation based on Hofmeister (2011) would appear more well-suited here.

Although this theory can be argued to have some merit, it also has considerable pitfalls. The greatest of these is the lack of clarity as to why Polish - and potentially other languages - exhibit this *which*-NP vs. *wh*-XP distinction for P-Stranding under sluicing, whereas others do not. Nykiel specifically puts forth the idea that 'preposition omission is attributable not to the syntactic form of a wh-remnant, but rather to an interaction of the semantic, syntactic, and phonological features of the remnant with the features of the correlate' (Nykiel, 2013, p.3), i.e. factors which have been documented to affect the ease of general anaphoric resolution. This would predict that all languages should conform to this preference pattern, something which does not appear to be the case (see, for instance, Serbo-Croatian (Stjepanović, 2008), Greek and German, here). Secondly, it is unclear why P-stranding is not rated as more acceptable than P-pied-piping, given it is closer to the ideal of Ariel's Accessibility Theory.

Another drawback of this theory, given it posits no structure at the *e*-site, is that it cannot explain the evident preference for case-matching between remnant and correlate. Sag & Nykiel (2011) attempt to tackle this by proposing an independently imposed sluicing prerequisite, whereby the sluice remnant must always match the correlate in number and case in order to be licensed. As we also saw, however, in section 1.2.5, such a restriction would fail to account for those instances where gender-, number- and/or case-mismatching is apparently also allowed under sluicing (see e.g. (40) and (41), repeated below). Moreover, it would not explain why a sprouted remnant must appear in a specific case, given there is no overt case-marked correlate to generate this restriction. For a more extensive series of arguments against such fragment theories, I will once again direct the reader to Merchant (2001).

(40) **Gender mismatching in Greek contrast sluicing**

O Petros sozi sinechos adespota zoa. Mechri
*Det.*M.NOM *Peter.*M.NOM *saves.*3SG *constantly* *stray.*N.ACC.PL *animals.*N.ACC.PL *until*

tora echi mazepsi dekapende skilous, alla den ksero poses
now *has.*3SG *collected* *fifteen* *dogs.*M.ACC.PL *but* NEG *know.*1SG *how.many.*F.ACC.PL

jates.
*cats.*F.ACC.PL

'Peter is constantly saving stray animals. Up until now he has saved fifteen dogs, but I don't know how many cats.'
[-1.8ex]

(41) **Number mismatching in English contrast sluicing**
Peter was on duty today at the airport, but I don't remember which other officers.

Theoretical issues aside, however, a more thorough examination of the studies presented in Nykiel (2013) also shows a number of design, methodological and analysis flaws, indicating that conclusions based on the results of these experiments should be drawn with a large degree of caution. For further examination of these experimental aspects, see section A.1 in the Appendix. The most important issues which I will draw the reader's attention to here, however, are, firstly, the uneven distribution of key characteristics across items, each of which could have plausibly affected the acceptability of P-stranding; and secondly, the lack of appropriate power for the type of analysis performed. With respect to the first point, this lack of feature control makes item comparisons problematic. Looking at an overview of the stimuli characteristics I present in table 1.1, based on an examination of the paper's Appendix, it is evident that there is an uneven number and distribution of a) prepositions, with 6 Ps presented an unequal number of times over 12 stimuli and one, furthermore, presented with two different cases; b) cases (Accusative, Genitive, Instrumental); and c) animate vs inanimate P complements. The verb types used were also inconsistent, mixing infinitives with participles, reflexives with non-reflexives, and di-transitives (with or without an additional PP) with monotransitives. The problem with not controlling for all these characteristics is that, once again, given the lack of available appropriately controlled research into each of these, it is impossible to know which ones could potentially be playing a significant role in the acceptability of a P-stranded remnant. For instance, it is unclear whether instrumental case appears significantly more frequently as an object with a P vs. without one (see Chapter 3 on why this might be a significant problem); given its use at times without a P almost as an adverbial phrase to express method (e.g. *Jadę autem* - 'I am going by car'; vs. *Jadę z autem* - 'I am going with the car' [e.g. to take it to be serviced]), it is unclear whether it may also be possible for a P-less Instrumental-marked remnant to be interpreted this way too. If this is true, then the Fit Condition would predict that a P-less remnant should also be acceptable, given it can easily slot into the full PP correlate's place in the antecedent without any further changes necessary.

Finally, there were, arguably, not enough data points per condition for a linear mixed effects model as used to give representative results. It is reported that 4 conditions and 12 experimental items were presented in the form of different questionnaires to 40 subjects. If we assume equal numbers of each condition were presented in each questionnaire version and across all questionnaires (in itself unclear), this would result in 12 items/4 conditions = 3 instances of each condition per subject; this translates to 40 (subjects) x 3 (instances per subject) = 120 instances per condition.

These issues, along with those reported in the Appendix, are evident in all of the paper's experiments. As such, although it may indeed be the case that P-stranding is allowed under sluicing in Polish and that it is governed by working memory mechanisms in its acceptability distribution, with *which*-NPs being preferred over *wh*-XPs, I do not consider

Table 1.1: Nykiel (2013) Exp. 1 Overview

			Number	Total
w (in/at/for)	Accusative	Inanimate	3	4
	Instrumental	Animate	1	
na (on/at/in/for)	Accusative	Inanimate	1	3
		Animate	2	
z (with/for)	Instrumental	Animate	1	1
za (behind/for/in/after/no translation)	Instrumental	Inanimate	1	1
od (from/since)	Genitive	Animate	2	2
o (about/at/for)	Instrumental	Inanimate	1	1

these experiments as clear evidence for or against either of these hypotheses.

1.3.4 Conclusion and the Cases of Greek and German

This brings us to the end of the various explanations that have been proposed thus far for what, at least at first glance, appear to be languages defying the PSG. A table with a more comprehensive list of languages together with their explanations is given in Table 1.2, at the end of this chapter.

As seen in the above sections and exemplified more clearly in the table, there are two major issues which plague the existing sluicing literature. The first of these concerns the generalisability of each explanation. Although situations with an acceptable alternative source are very convincing, these can only apply to those languages where such a source is available, i.e. predominantly non-case-marking languages for a cleft source, and languages where resumption can indeed salvage P-stranding for a resumptive pronoun source. This does not apply to all PSG-defying languages, however, and all other explanations simultaneously run into their own share of obstacles to varying degrees of severity. The major recurring issue and common theme to these obstacles is that these explanations fail to convincingly generalise to other, similar languages, either by over-generalising and predicting PSG-defiance where none has been documented (e.g. PPs should technically always be PF-islands if they are for some languages), or, conversely, failing to account for this defiance where it has been - arguably - 'documented' (e.g. P-stranding being acceptable despite the barring factor of DP-coalescence, as would be predicted by Sato (2011)).

This last point, in particular, brings me to the next key issue plaguing the entirety of the existing sluicing literature with respect to any language, and that is the lack of adequate experimental datasets. With the exception of AlShaalaan & Abels (2019), none of these papers contain data collected from adequately powered and well-controlled experiments.

Instead, the majority contain only anecdotal evidence based on the linguistic judgements of one or two native speakers, predominantly those of the researchers themselves. That the judgements of linguists too closely related to the subject at hand are not always representative of those of naïve native speakers is something all linguists should be intimately aware of, however the reader is also referred to Dąbrowska (2010) for more evidence of this. On the other hand, what little experimental data *is* reported is too small a dataset per study to be considered representative or accurate for the analyses conducted, with each study, furthermore, containing multiple confounding factors. As a result of these issues, although both connectivity effects have been widely theoretically accepted in the literature, in reality we have little to no idea what they look like in terms of acceptability and whether they indeed align with the linguistic judgements of naïve native speakers. Is P-stranding truly as black and white in acceptability as the theoretical literature would argue it is? Or are we in fact dealing with more of a grey, gradient area of acceptability, which would, in turn, require further explanation beyond what has thus far been considered?

This thesis attempts to tackle both of these issues. On the one hand, I will address the lack of experimental evidence by contributing several novel, large-scale and thoroughly controlled experimental datasets from two overtly non-P-stranding languages, namely Modern Greek and German. In all the existing theoretical literature, these two languages have been proposed multiple times to be especially robust examples of both form-identity generalisations. In Merchant (2001), thanks to the author's in-depth knowledge of the two languages and their rich feature-marking, it is not an overstatement to say that Greek and German function as the cornerstones of these generalisations, with over a dozen Greek examples and close to 40 standard and Swiss German examples used to demonstrate their validity (a few of which are presented below). As a result, both languages go on to be quoted several more times in various papers as compliant with the PSG, for instance in Van Craenenbroeck (2009, 2010a), a.o. What all these papers have in common, however, as expected, is a lack of any experimental investigation on the acceptability of the examples used. When practically put to the test, I will show that both of these languages do in fact allow P-stranding under sluicing to a significantly higher degree of acceptability than they do overtly, albeit not to the same degree as they do P-pied-piping.

Furthermore, I will show that no previously proposed explanation, whether based on syntactic theory or sentence processing, can account for these results. Instead, in order to explain both these data and the previously proposed behaviour of other languages, I will be arguing for a novel approach which integrates theoretical syntax with sentence - and, more generally, information - processing. Specifically, I believe that combining a theoretical account which posits structure at the *e*-site with a Bayesian estimation approach which treats sluicing as an inherently 'noisy' communication channel can best capture both these and previous cross-linguistic data. In this way, I intend to address the other major problem identified for the existing sluicing literature, i.e. the lack of a cohesive and cross-linguistically generalisable explanation. Based on this novel approach, I argue that all overtly non-P-stranding languages should show similar acceptability results for this phenomenon when practically tested, with no such thing as black and white acceptability or unacceptability of P-stranding under sluicing, but instead a broader, grey area. I shall argue that what affects the nature of this acceptability result within this grey area for each language is how easily the parser can infer a well-formed likely structure within the *e*-site. The harder such a structure is to reach, the higher the processing cost will be to parse the utterance and, by extension, the lower its acceptability rating. Conversely, the easier it is to reach, the easier the utterance will be to parse, and, by extension, the higher its acceptability rating. In this way, simple and easily testable cross-linguistic predictions are generated, forgoing the need to create different, elaborate and idiosyncratic theoretical explanations for each language individually. I also provide a thorough experimental methodology for testing these predictions and outline several additional experiments which can be used to further confirm or disconfirm this new appraoch. Finally, regardless of whether or not this account is correct, I argue that such predictions and ways of proving or disproving them *must* be made across all languages if this - or any other - theory is to have any merit, with the constant creation of new rules per language examined being a logistically inefficient way forward, not to mention its implausibility from a learnability perspective.

Table 1.2: PSG-defying Languages, alphabetically

Language	Authors	Explanation	Drawbacks
Bahasa Indonesian	Fortin, 2007	LF-Copying	1. No experimental data; 2. No explanation for other types of languages
	Sato, 2011	PF-Interface Repair, based on [+wh] feature percolation	1. No experimental data; 2. Inadequate explanation for other types of languages based on P+DP coalescense.
Brazilian Portuguese	Almeida & Yoshida, 2007	PP as PF-Island	1. No experimental data; 2. Unclear why PPs are PF-islands for some, but not other languages
	Rodrigues et al., 2009	Cleft source	1. No experimental data; 2. No explanation for other types of languages
Emirati Arabic	Leung, 2014	PP as PF-Island	1. No experimental data; 2. Unclear why PPs are PF-islands for some, but not other languages
Jordanian Arabic	Al Bukhari, 2016	Cleft source + Resumption	1. No experimental data; 2. No explanation for other types of languages
Polish	Szczegielniak, 2008	Cleft source	1. No experimental data; 2. Cleft source shown to be impossible (Nykiel, 2013)
	Nykiel, 2013	Pointer mechanism	1. Inadequate experimental data; 2. Independently imposed case-matching requirement overgenerates.
Saudi Arabic	AlShaalan & Abels, 2019	Resumption	1. Good experimental data, but no explanation for results when resumption blocked; 2. No explanation for other types of languages.
Serbo-Croatian	Stjepanovic, 2008	No explanation	No explanation
Spanish	Rodrigues et al., 2009; Vicente, 2008	Cleft source	1. No experimental data; 2. No explanation for other types of languages

Chapter 2

Cross-linguistic Acceptability Judgement Data

2.1 Experiment 1: P-Stranding under Regular Sluicing

2.1.1 Experimental Outline, Aim and Predictions

At the end of Chapter 1, it became evident that the two predominant issues besetting the sluicing literature are its lack of a cohesive explanation for apparent PSG-defying behaviour cross-linguistically, as well as a lack of adequate experimental evidence for this anecdotally reported behaviour. This chapter shall address the latter problem, in particular, by presenting four large-scale acceptability judgement studies targeting Greek and German P-stranding.

The first experiment presented investigates the behaviour of Greek with respect to the two form-identity generalisations, both overtly and under sluicing. Given these two generalisations have been proposed to form a single class of behaviour under sluicing, matching their overt counterparts with respect to syntactic rule conformation in disallowing both P-stranding and case-mismatching (in opposition to islands, for the most part), although this first experiment was designed to primarily check the P-Stranding Generalisation in Greek, it made sense to take this opportunity to also secondarily confirm this prediction of identical behaviour regrading the Case-Matching Generalisation. If P-stranding and case-mismatching are significantly less acceptable than P-pied-piping and case-matching overtly, as syntactic theory claims, then the goal of this study is to identify how the presence of sluicing may affect the size of this difference.

As such, to investigate both form-identity generalisations and whether indeed they present as a single behaviour class under sluicing, Experiment 1 crosses the factors of Sluicing[1] (±, i.e. sluicing vs. overt continuation of sluice), P-stranding (±, i.e. P-stranding vs. P-pied-piping), and Case-matching (±, i.e. case-matching of remnant with correlate or case-mismatching), resulting in the table shown in 2.1[2]. In this way, it compares how Greek speakers judge P-stranding and case-mismatching compared to P-pied-piping and case-matching, both overtly and under sluicing, allowing us to identify a) a baseline for how acceptable or unacceptable P-stranding and case-mismatching are overtly compared to P-pied-piping and case-matching; b) whether and how these differences might change under sluicing. Although we are predominantly interested in the interactions of P-Stranding with Sluicing, on the one hand, and Case-matching with Sluicing, on the other hand, this fully crossed design will also allow us to investigate whether P-stranding and case-mismatching may together have an additive effect on a sentence's acceptability overtly and/or under sluicing. The experiment examines each sentence's acceptability by using a tightly controlled large-scale web-based acceptability judgement study utilising a 7-point Likert-type rating scale. The study was also designed to check whether any of the previously proposed explanations for apparent PSG-defying languages may explain any such Greek finding, as shall be made clearer when the stimulus design is explained in more detail in section 2.1.2.4[3].

In terms of predictions for these factors, let us begin with the most straightforward and unequivocal ones. First and foremost, regardless of the theoretical approach one entertains, Greek is widely accepted to be an overtly non-P-stranding language. All theories should, therefore, agree in their prediction for a significant difference between P-stranding and P-pied-piping at least in overt conditions with P-pied-piping being predicted to be always significantly more acceptable than P-stranding overtly. Secondly, regardless of theoretical approach, one would also anticipate that there should be a significant difference between case-matching and case-mismatching, at least overtly, with case-mismatching being considered significantly less acceptable than case-matching. Whether or not these results should present as main effects or be tempered by an interaction between factors, that is where the various approaches begin to differ in their expectations.

Specifically, with respect to the factor of P-Stranding, when we do not take into consideration Case-matching, a PF-Deletion Hypothesis would predict judgements of P-stranding conditions to be uniformly unacceptable, both overtly and under sluicing. This should be in polar opposition to their P-pied-piped counter-parts; in other words, a PF-Deletion approach

[1] From here on, capitalisation within the context of experiments will be used to signify factors, with small letters signifying levels of the factor, e.g. 'Sluicing' vs. 'sluicing' and 'P-Stranding' vs. 'P-stranding'.

[2] This study, along with almost all subsequently listed studies, was approved by the UCL Research Ethics committee under the Project ID Number *LING-2012-3*, with Dr Andrea Santi as primary investigator. The German acceptability study, in particular, was approved under the project ID Number *LING-2015-03-23*, with Dr Klaus Abels as primary investigator. All subjects read the UCL Information and Ethics form prior to beginning the experiment and expressly agreed to participate knowing they were free to withdraw at any point. All data collected prior to May 2018 were handled according the the 1998 Data Protection Act, and all data collected subsequently were handled according the 2018 EU General Data Protection Regulations.

[3] This study is a more tightly controlled replication of a study I conducted as an MSc project at UCL for an advanced syntax course term paper.

Table 2.1: Experiment 1 Conditions and Examples

(55) Stin avli i neari mathitria krivotan
In.Det.F.ACC *school.yard*.F.ACC *Det*.F.NOM *young*.F.NOM *student*.F.NOM *was.hiding*.3SG
apo kapjous, alla kanis den katalave...
from someone.M.ACC *but no.one*.M.NOM NEG *understood*.3SG...
'In the school yard, the young student was hiding from some people, but no one understood...'

	Condition	Example
1	Case-Matching, Non-Sluicing, P-Pied-Piping	... apo pjous krivotan. ... *from who*.M.ACC *was.hiding*.3SG '... from who she was hiding.'
2	Case-Matching, Non-Sluicing, P-Stranding	... pjous krivotan apo. ... *who*.M.ACC *was.hiding*.3SG *from* '... who she was hiding from.'
3	Case-Matching, Sluicing, P-Pied-Piping	... apo pjous. ... *from who*.M.ACC '... from who.'
4	Case-Matching, Sluicing, P-Stranding	... pjous. ... *who*.M.ACC '... who.'
5	Case-Mismatching, Non-Sluicing, P-Pied-Piping	... apo pji krivotan. ... *from who*.M.NOM *was.hiding*.3SG '... from who she was hiding.'
6	Case-Mismatching, Non-Sluicing, P-Stranding	... pji krivotan apo. ... *who*.M.NOM *was.hiding*.3SG *from* '... who she was hiding from.'
7	Case-Mismatching, Sluicing, P-Pied-Piping	... apo pji. ... *from who*.M.NOM '... from who.'
8	Case-Mismatching, Sluicing, P-Stranding	... pji. ... *who*.M.NOM '... who.'

would predict a significant difference between P-stranding and P-pied-piping conditions, regardless of Case-matching, a difference which should remain unaffected by the presence or not of sluicing. This would translate to a main effect for the factor of P-Stranding, with no significant interaction with that of Sluicing. The exact same pattern of behaviour would also be predicted to be manifest with regards to the factors of Case-matching and Sluicing, when isolated from P-Stranding manipulations, i.e. when we collapse across P-pied-piping conditions. Specifically, such case-mismatched conditions should always be deemed unacceptable compared to case-matching conditions, with a significant difference predicted between the two. This difference should, furthermore, be unaffected by the factor of Sluicing. In other words, we would anticipate a significant main effect of Case-matching with no significant interaction of Case-matching with Sluicing in any direction. This lack of interaction between Sluicing and P-Stranding, on the one hand, and Sluicing and Case-matching, on the other hand, can be encapsulated by the idea that sluicing should not make P-stranding and/or case-mismatching any more or less acceptable than they are overtly. As a result, we would not anticipate a significant difference between overt and sluicing conditions, all else being equal, i.e. we would not anticipate a main effect for the factor of Sluicing.

In the above predictions we have examined P-Stranding and Case-matching separately from each other, i.e. manipulating each factor and examining its predicted interaction with the factor of Sluicing whilst collapsing across the other factor. Even when we examine all three factors together, however, given its lack of gradient acceptability predictions, things do not change greatly from the point of view of a PF-Deletion Hypothesis. Specifically, based on everything we have seen thus far, it is clear that such a hypothesis would predict that both case-mismatching and P-stranding should always be at floor-level in terms of acceptability; this means that regardless of the context they appear in, i.e. regardless of whether we are talking about an overt continuation or a sluice, it should be impossible to drive acceptability any further down (or up, for that matter) by combining P-stranding together with case-mismatching or vice versa. As such, a PF-Deletion Hypothesis would not anticipate any factor (super)additivity effects or, in other words, a significant three-way interaction between P-stranding, Case-matching and Sluicing in any direction.

Moving on to the other accounts we saw in Chapter 1, an LF-Copying account would make similar predictions with respect to the isolated[4] factors of Case-matching and Sluicing. That is to say, thanks to the stipulated structural isomorphism requirement between correlate and remnant, it would predict that case-mismatching should always be significantly worse than case-matching, regardless of the presence or not of sluicing. This would translate to a main effect of Case-matching with no predicted interaction with the factor of Sluicing in any direction. That having been said, when we collapse across Case-matching, an LF-Copying approach would not be able to clearly make a prediction with respect to

[4]It should be made clear that I use the term 'isolated' rather loosely here to indicate that two factors are being manipulated at this point whilst keeping the third factor 'stable' by collapsing across it.

P-stranding under sluicing unless the structural isomorphism it calls for between remnant and correlate extends not only to case, but also to phrasal category. The issue here, of course, as also mentioned in Chapter 1 would be that such an extension would force even overtly P-stranding languages to always have P-pied-piped sluice remnants if the antecedent also contains P-pied-piping. In other words, examples such as (22), repeated below, should be considered unacceptable in English, which is evidently not the case.

(22) Mary danced with someone, but I can't remember who.

If we were to make this - rather unlikely - extension, then we would also anticipate the same predictions regarding the factors of P-Stranding and Sluicing as a PF-Deletion account, i.e. when isolated from Case-matching, P-stranding should always be significantly worse than P-pied-piping, with the factor of Sluicing not having a significant effect on the size of this difference. Similarly to a PF-Deletion account, an LF-Copying account does not make any allowances or predictions for (super)additivity effects, with 'unacceptable' conditions all being predicted to receive the minimum allowable ratings and P-stranding not making case-mismatching any better or worse or vice versa, regardless of the presence or not of sluicing. As such, as in the case of a PF-Deletion Hypothesis, we would not predict significant two- or three-way interactions between Case-matching, P-stranding and/or Sluicing. Overall, given sluicing should in no way impact upon the acceptability of sentences, there should also, therefore, be no significant difference between overt and sluicing conditions, all else being equal, i.e. we would not anticipate a main effect for the factor of Sluicing.

A pointer account, on the other hand, would make rather different predictions. In general, eschewing structure at the *e*-site would mean that both P-stranding and case-mismatching should instantly become acceptable under sluicing, even if they are not overtly. In other words, even though P-stranding and case-mismatching may be significantly different to P-pied-piping and case-matching, respectively, in overt conditions, a pointer account would predict that these differences should be minimised to zero under sluicing, with all conditions being similarly acceptable and at ceiling level. As such, a significant two-way interaction would be predicted between the factors of P-Stranding and Sluicing, on the one hand, whilst keeping Case-matching stable, and Case-matching and Sluicing, on the other, whilst keeping P-Stranding stable. However, even though all sluicing conditions should be considered equally acceptable, given their lack of structure, it is not clear exactly how a pointer account would handle P-stranding together with case-mismatching overtly. In other words, although this is a processing-based account with respect to how sluicing itself is handled, overt structures are not explicitly dealt with. Given overt structures are left to the domain of overt syntactic theory, it would thus be reasonable to assume that a pointer account would make the same predictions as an LF-copying or PF-Deletion account with respect to P-Stranding and Case-matching in the context of non-sluicing continuations, i.e.

when combined, P-stranding and case-mismatching should not have a compounding effect, instead being judged at floor level and equally to case-matching with P-stranding overtly and case-mismatching with P-pied-piping overtly. As such, given overtly P-stranding and case-mismatching should always be judged at floor level, whether combined with each other or not, and given under sluicing P-stranding and case-mismatching should also always be judged at ceiling level, whether combined or not, we would thus not expect a three-way interaction between factors in any direction. Finally, given overt conditions should exhibit a variety of acceptability ratings based on the presence or not of case-mismatching and P-stranding, whereas all sluicing conditions would be expected to be equally acceptable and at ceiling level, a pointer account should also predict a main effect for the factor of Sluicing, with sluicing conditions being overall significantly more acceptable than overt ones.

This is one way of thinking about the pointer account. If, however, we adopt the view of a pointer account in the style of Sag and Nykiel (2011) and Nykiel (2013), then these predictions need to be slightly altered. Specifically, with respect to the factor of Case-matching firstly, setting aside P-Stranding for the moment, by imposing a case-matching requirement between correlate and remnant, such an account would anticipate case-mismatching conditions to be uniformly unacceptable and at floor level, both overtly and under sluicing, performing significantly worse than their case-matching counterparts. In other words, the presence or not of sluicing should in no way impact upon the acceptability of case-mismatching, i.e. there should be no two-way interaction in any direction between Case-matching and Sluicing, with a significant main effect of Case-matching predicted instead. With respect to the factor of P-Stranding, leaving Case-matching aside for the moment, the account's predictions for overt conditions are no different to those of any of the previous accounts. However, under sluicing, this version of the pointer account requires that we take into consideration how the complexity of the remnant and/or correlate may be a significant predictor of the acceptability of P-stranded remnants (Nykiel, 2013). In other words, we cannot make a simple prediction about a two-way interaction between P-Stranding and Sluicing, instead predicting that the difference between P-stranding and P-pied-piping under sluicing be significantly modulated by the additional factor of (remnant) Complexity, such that only *which*-NP remnants, referring to more complex correlates, should allow P-stranding. When examining the factors of Case-matching and P-Stranding together, as far as overt conditions are concerned, this account would make the same predictions as all previous ones, with case-mismatching and P-stranding being always judged at floor level and with no (super)additivity effects. Under sluicing, case-mismatching conditions should also always be at floor level, regardless of the presence or not of P-stranding; for case-matching conditions, P-stranding should be acceptable only if we are dealing with a complex *which*-NP remnant. As a result, this account would not anticipate a three-way interaction between Sluicing, P-Stranding and Case-matching, but rather a significant three-way interaction between Sluicing, P-Stranding and Remnant Complexity. Given these particular predictions can become a little complex with respect to factor interactions, the reader is referred to acceptability predictions per condition presented in Table 2.2 for this pointer account in particular.

Table 2.2: Nykiel (2013) Pointer Account Acceptability Predictions

Non-sluicing	Case-matching	P-pied-piping	which-NP	✓
			who	✓
		P-stranding	which-NP	✗
			who	✗
	Case-mismatching	P-pied-piping	which-NP	✗
			who	✗
		P-stranding	which-NP	✗
			who	✗
Sluicing	Case-matching	P-pied-piping	which-NP	✓
			who	✓
		P-stranding	which-NP	✓
			who	✗
	Case-mismatching	P-pied-piping	which-NP	✗
			who	✗
		P-stranding	which-NP	✗
			who	✗

In the interest of clarity, a table reflecting the predictions of each of these major theories is presented in Table 2.3.

Table 2.3: Experiment 1 Significance Predictions

	Prediction of Significance	PF-Deletion	LF-Copying	Pointer	Pointer (Nykiel)
1	Sluicing	✗	✗	✓	✗
2	Case-Matching	✓	✓	✗	✓
3	P-Stranding	✓	✗	✗	✗
4	Sluicing*Case-Matching	✗	✗	✓	✗
5	Sluicing*P-Stranding	✗	?/✗	✓	✓
6	Sluicing*P-Stranding*Case-Matching	✗	✗	✗	✗
7	Sluicing*P-Stranding*Complexity	✗	✗	✗	✓

2.1.2 Methods

2.1.2.1 Experimental Design

The three binary factors of Sluicing, P-Stranding and Case-matching were crossed to yield a total of eight (8) minimally different experimental conditions (Table 2.1, repeated below)[5]. All conditions began with the structure shown in (55) and continued as shown in the Table.

table below + refer-
ng

These conditions were presented in a Latin Square design to participants, such that each participant saw only one condition per item, and equal numbers of each condition overall. In this way, a total of 64 items were presented to each participant, with 8 instances of each condition. In addition, twice as many fillers (128) as items were interspersed amongst these in order to keep participants naïve as to the true purpose of the experiment, with comprehension questions also used to test participants' attention. Each of these aspects is described in more detail below.

2.1.2.2 Method

A web-based acceptability judgement task hosted on Alex Drummond's Ibex Farm® platform (www.spellout.net/ibexfarm/) was used to measure participants' linguistic intuition regarding these conditions. Following an extensive literature review on the use of rating scales in psychology since Hayes and Patterson (1921) and Freyd (1923), investigating whether to use analog or radio buttons, individual labelled (or non-labelled) rating points along a continuum or not, and indeed how many such rating points there should be, a bipolar scale of 7 points was found to have the greatest cross-sectional reliability (Birkett, 1986; Komorita & Graham, 1965), i.e. reliability in obtaining similar ratings from a given person for the same attitude at any one time (Krosnick & Fabrigar, 1997), but also longitudinal reliability, i.e. reliability in obtaining similar ratings from a given person on the same question after a specific period of time (Alwin & Krosnick, 1991).

I hope the reader will allow me a small digression here to explain why this method was specifically selected over another, quite widely cited method of obtaining subjective judgements, the Magnitude Estimation (ME) task, which has been proposed to be more accurate than a Likert-scale. This method, originally proposed for linguistic judgements by Bard, Robertson, and Sorace (1996) originated in the field of psychophysics to more accurately capture participants' perception of physical stimuli and to overcome potential scaling issues that *n*-level rating scales might introduce. Specifically, instead of using a binary or *n*-level rating scale, ME allows participants to directly compare perceived stimuli to each other by using a specific stimulus as a measure or point of reference. E.g. if stimulus

[5]Non-sluiced conditions were created using a continuation of the sluiced ones in order to keep them minimally different and not introduce new material.

Table 2.4: Experiment 1 Conditions and Examples

(55) Stin avli i neari mathitria krivotan
In.Det.F.ACC *school.yard*.F.ACC *Det*.F.NOM *young*.F.NOM *student*.F.NOM *was.hiding*.3SG
apo kapjous, alla kanis den katalave...
from someone.M.ACC *but no.one*.M.NOM NEG *understood*.3SG...
'In the school yard, the young student was hiding from some people, but no one understood...'

	Condition	Example
1	Case-Matching, Non-Sluicing, P-Pied-Piping	... apo pjous krivotan. ... *from who*.M.ACC *was.hiding*.3SG '... from who she was hiding.'
2	Case-Matching, Non-Sluicing, P-Stranding	... pjous krivotan apo. ... *who*.M.ACC *was.hiding*.3SG *from* '... who she was hiding from.'
3	Case-Matching, Sluicing, P-Pied-Piping	... apo pjous. ... *from who*.M.ACC '... from who.'
4	Case-Matching, Sluicing, P-Stranding	... pjous. ... *who*.M.ACC '... who.'
5	Case-Mismatching, Non-Sluicing, P-Pied-Piping	... apo pji krivotan. ... *from who*.M.NOM *was.hiding*.3SG '... from who she was hiding.'
6	Case-Mismatching, Non-Sluicing, P-Stranding	... pji krivotan apo. ... *who*.M.NOM *was.hiding*.3SG *from* '... who she was hiding from.'
7	Case-Mismatching, Sluicing, P-Pied-Piping	... apo pji. ... *from who*.M.NOM '... from who.'
8	Case-Mismatching, Sluicing, P-Stranding	... pji. ... *who*.M.NOM '... who.'

x is loud, then stimulus y could be judged twice as loud as x, half as loud etc. Although this method can be argued to have some advantages over Likert-type scales, such as subjects having more flexibility in assigning a judgement by being able to choose just how finely grained they would personally like the scale to be, I believe a number of disadvantages make it less than ideal. The most important advantage attributed to this method is that participants would be able to use it to rate sensations on a ratio scale, as opposed to an interval scale, as with a Likert one. That is to say, differences in sensations are perceived as multiples of each other, as mentioned above. This brings with it two caveats. Firstly, for these ratio estimations to make sense, basic algebra dictates that an absolute zero point is required, i.e. a point below which it is impossible to go. When this scale was being used to measure other sensations in its original field of psychophysics, such as taste, smell, pain or hearing, one could potentially argue for such an absolute zero point: when a subject is eating or smelling nothing, then they are not technically tasting anything, hence this can serve as an absolute zero point in their sense of taste. In terms of linguistic judgements, however, a point of absolute zero acceptability is much less clearly definable, particularly when required to be generaliseable across all native speakers. Even the use of a completely garbled sentence or perhaps a sentence in a non-linguistic form would not be useable in a way that makes other sentence judgements comparatively meaningful. The second caveat is that people must indeed have the cognitive capacity to make judgements on such a ratio scale. Since ME was first introduced, this capacity has been argued to in fact not exist in humans (see Luce (2002) and Narens (1996) for more on the necessary assumptions of this capacity, as well as Sprouse (2011) for how they do not hold for linguistic judgements, in particular). Furthermore, what have been argued to be ME's largest advantages, i.e. its theoretical additional sensitivity above and beyond the finite n-levels of a defined scale, and how it can overcome scaling issues which can theoretically plague n-level rating scales, do not actually appear to be practically true. Studies have found no apparent sensitivity advantage of ME over more traditional acceptability judgement scales (Bader & Häussler, 2010; Myers, 2009), with some evidence that it can, in fact, introduce unrepresentative variance in scores (Weskott & Fanselow, 2011). Based on these issues, we adopt the stance here taken by other experimental linguists such as Featherston (2008); Sprouse (2007, 2011); Weskott and Fanselow (2011) that ME is not appropriate for judging linguistic acceptability. In all acceptability judgement studies presented here, therefore, a clearly outlined 7-point scale was used, with 1 being the lowest possible rating (completely unacceptable) and 7 the highest (completely acceptable).

Each sentence was presented in full in the centre of the screen, with the detailed 7-point scale beneath it. Participants were instructed to read the sentence at a natural pace and select what they deemed an appropriate rating using the keyboard or mouse. In order to minimise the possibility of participants being distracted by other tasks whilst reading the sentences, a timer of 45 seconds was set, after which the program would move on to the next screen, recording a non-response. After 2/3 of unequivocally acceptable experimental sentences, i.e. P-pied-piped case-matching conditions, and half of the grammatical fillers, a non-timed yes/no comprehension question was presented. Between stimuli, a non-timed pause screen with central fixation cross was presented, with participants pressing a key to continue, allowing them to take multiple breaks wherever needed.

Detailed instructions and examples were provided prior to the experiment, along with 3 practise trials. All examples and practise trials contained manipulations unrelated to those of the experiment. Given the large number of participants involved in the study and that appropriate compensation for each was not financially feasible at the time, participants were each entered in a £100 prize draw upon completion of the study. Once the study was over, one entrant was randomly selected and awarded the prize.

2.1.2.3 Participants

A total of 125 native Greek speakers from across Greece were recruited through Facebook® over the course of two weeks. Their native speaker status was judged based on their own word, but also a short paragraph (approximately 8 sentences) which they were asked to write on any subject they wished before taking part in the experiment. Out of 125 participants, the data of only 84 were used. Although this may seem like a large number of excluded participants, one must take into account the non-controllable experimental environment of such web-based studies, where subjects can and do often get distracted without the experimenter's knowledge, resulting in a much higher exclusion rate than in more controlled, face-to-face studies. The criteria for participant exclusion were a) less than 80% accuracy on all comprehension questions; b) rating all conditions the same or inverting the rating scale despite the clearly detailed scaling scheme beneath each sentence; c) writing an incomprehensible or error-filled paragraph at the beginning of the experiment; or d) a combination of the above.

2.1.2.4 Items

2.1.2.4.1 Experimental Stimuli

Stimuli were created following the basic pattern delineated in (56), for instance (57).

(56) Main Clause Subject + Main Verb + Preposition + Indefinite Internal Argument, 'but' (+ Second Clause Subject) + NEGATION + Embedding Verb + Interrogative Pronoun (Accusative).

(57)	I	Maria	chorepse	me	kapjon,	alla	den	ksero	pjon.
	*Det.*F.NOM	*Maria.*F.NOM	*danced.*3SG	*with*	*someone.*M.ACC	*but*	NEG	*know.*1SG	*who.*M.ACC

'Maria danced with someone, but I don't know who.'

This pattern was obviously adapted wherever necessary to conform to the experimental conditions, but also embellished in the main clause with non-discourse-related material, such as adverb phrases, to create more natural-sounding sentences. The stimuli were, furthermore, subject to a strict set of rules to control for various arguments which have been - or may be - proposed to act as confounding factors influencing the acceptability of P-stranding. First of all, the main verb (e.g. *danced*) was chosen to only select for PP complements and not bare Accusative-marked complements. This was so that the bare Accusative-marked remnant could not possibly be perceived as a direct or indirect object of the main verb, with or without the same meaning. If such an interpretation were allowed, then this would obviously have made the remnant much more acceptable independently of the effects of P-stranding. Although this is a basic and fundamentally necessary requirement, previous experiments on P-stranding under sluicing do not appear to have taken it into consideration (see e.g. Nykiel (2013)). Secondly, with respect to the P's complement, only masculine NPs were used due to their overt feature-marking suffixes which clearly distinguish between Nominative and Accusative case in both singular and plural[6]. This was done so that the remnant could not possibly be interpreted as the pivot of a cleft or copular structure (e.g. as in Rodrigues et al. (2009) for BP), since cleft pivots must always appear in Nominative in Greek. If the two forms were confounded, this would potentially allow such an alternative, non-P-stranding source to be inferred. Thirdly, with respect to the main clause's subject, only feminine, clearly gender-marked NPs were used to avoid the potential for competition interference based on shared morphosyntactic features during the parser's attempt to find the correlate within the main clause (Caramazza, Grober, Garvey, & Yates, 1977; Criss & McClelland, 2006; Ehrlich, 1980; Hofmeister, Jaeger, Arnon, Sag, & Snider, 2013; Nairne, 1990, 2001, 2006; Oberauer & Lewandowsky, 2008). Furthermore, with respect to the form of the *wh*-remnant itself, half of these were simple interrogative pronouns (*who*) and half were *which*-NPs, with an overtly expressed NP. The simple interrogative pronouns referred back to a simplex correlate (*someone*) and the *which*-NPs referred back to a complex one (*some bank account*). This particular manipulation was included to address Nykiel (2013)'s claim that the form of the remnant and/or the correlate, in terms of complexity, may affect how easy the correlate is to retrieve and, by extension, how easy it is to P-strand the remnant. Moreover, for the case-mismatching conditions, all mismatches were created using Nominative-marking instead of Accusative. With regard to the number of the remnant and correlate, half the stimuli included a singular-numbered correlate/remnant and the other half a plural one. In Greek feature-marking, the orthographic string difference between Nominative and Accusative is smaller for singular numbering (*ποιος/pjo-s* vs. *ποιον/pjo-*

[6]Such case-distinct feature-marking is apparent only in masculine nouns, certain, relatively uncommon feminine nouns of archaic declension and not at all in neuter nouns.

n), than for plural numbering ($\pi o\iota o\iota$/*pj-i* vs. $\pi o\iota o\upsilon\varsigma$/*pj-us*). I decided to include this factor because in the original MSc study this experiment was based on, only singular-marking was used, and the results of that study had shown a slight increase in acceptability for case-mismatching under sluicing. One explanation for this increase in acceptability is that participants were potentially misreading the sentence-final letter, misinterpreting the Nominative for Accusative, i.e. a case-matching remnant. As a result, I decided to include plural-marking in this study to check for this possibility, since participants should be less likely to misread when more letters are involved, particularly when dealing with a plural *which*-NP involving multiple letter differences between Nominative and Accusative across two words. With respect to the Ps themselves, equal numbers of the four most common ones were used: 'me' (*with*), 'se' (*to*), 'jia' (*for*), 'apo' (*from*).

One final factor which was controlled for is related to the concept that apparently acceptable P-stranding under sluicing could, in fact, be due to a temporary grammatical illusion (see section 2.6.2 for more details). Specifically, it could be argued that what is intended as a P-stranded remnant is, in fact, being (temporarily) interpreted as the embedding verb's direct object (e.g. as the object of *know*), thereby generating the illusion of acceptability for what should technically be considered unacceptable. To check for this possibility, I designed the stimuli so that in half of them the embedding verbs could *c*-select for an Accusative-marked NP complement, and in the other half they could only *c*-select for a Genitive-marked NP or a full CP.

All of the above mentioned factors were controlled for and equally distributed across items to yield the table in 2.5. In addition to the above design constraints, the stimuli used were further controlled for sentence length, word frequency and plausibility[7].

2.1.2.4.2 Sentence Length

All sluice conditions were measured and controlled for sentence length. Non-sluice conditions were not directly compared to these as they necessarily varied quite drastically. Descriptive statistics showed sluice sentence lengths to be normally distributed (Mean: 19.03; Median: 19.00; SD: 2.20; Shapiro-Wilk: $p = .094$, *n.s.*), as were those of fillers (Mean: 18.21; Median: 18; SD: 2.25; Shapiro-Wilk: $p = .098$, *n.s.*).

2.1.2.4.3 Word Frequency

Although word frequency is very rarely controlled for in acceptability judgement tests in the literature, there appear to be divided opinion and experimental results as to its actual effects. On the one hand, emergentists consider structure to emerge as a result of natural

[7]In order to reduce inter-experiment variability, significant effort was made to use as similar items as possible between experiments. As such, this particular section will be referred back to when the next experiments are presented.

Table 2.5: Experiment 1 Stimuli Organisation

<table>
<tr><th>Preposition</th><th>Remnant Form</th><th>Number</th><th colspan="3">Total</th></tr>
<tr><td rowspan="4">me</td><td rowspan="2">who</td><td>Singular</td><td>4</td><td rowspan="4">16</td><td rowspan="16">64</td></tr>
<tr><td>Plural</td><td>4</td></tr>
<tr><td rowspan="2">which</td><td>Singular</td><td>4</td></tr>
<tr><td>Plural</td><td>4</td></tr>
<tr><td rowspan="4">se</td><td rowspan="2">who</td><td>Singular</td><td>4</td><td rowspan="4">16</td></tr>
<tr><td>Plural</td><td>4</td></tr>
<tr><td rowspan="2">which</td><td>Singular</td><td>4</td></tr>
<tr><td>Plural</td><td>4</td></tr>
<tr><td rowspan="4">jia</td><td rowspan="2">who</td><td>Singular</td><td>4</td><td rowspan="4">16</td></tr>
<tr><td>Plural</td><td>4</td></tr>
<tr><td rowspan="2">which</td><td>Singular</td><td>4</td></tr>
<tr><td>Plural</td><td>4</td></tr>
<tr><td rowspan="4">apo</td><td rowspan="2">who</td><td>Singular</td><td>4</td><td rowspan="4">16</td></tr>
<tr><td>Plural</td><td>4</td></tr>
<tr><td rowspan="2">which</td><td>Singular</td><td>4</td></tr>
<tr><td>Plural</td><td>4</td></tr>
</table>

language use and therefore to be highly influenced by frequency; their concept of 'entrenchment' refers to our repeated exposure to specific forms and structures cementing them as the basis of our linguistic constructs (Ambridge, Pine, Rowland, & Young, 2008); their complementary concept of statistical preemption refers to repeated usage of certain structures over others rendering the latter more obsolete and thus unacceptable to speakers. On the other hand, nativists take the stance that usage frequency does not influence the acceptability of structures at all, given these are 'inherent' or 'native' to a language (Bermel & Knittl, 2012). Experimentally, some studies have shown frequency to have an effect on acceptability (Bermel & Knittl, 2012), whereas others have not (Divjak, 2008). In the interest of being as accurate as possible, all the Greek studies presented here have been controlled for word frequency. This has the additional advantage of opening up the possibility for these stimuli to be used in their current form in future both self-paced reading and ERP studies, given the well-established effect of this factor on reading times and in EEG experiments (Embick, Hackl, Schaeffer, Kelepir, & Marantz, 2001; Van Petten & Kutas, 1990).

Word frequency was assessed based on the SUBTLEX-GR[8] files of the 2010 study by Dimitropoulou, Duñabeitia, Avilés, Corral, and Carreiras. Excluding all functional words and proper names, all words contained within each item were assessed separately for word frequency and then the average word frequency calculated per item. All words used had a Log10 frequency per million (fpm) value larger than 1.5, given word frequency effects have been argued to be largest below a Log10 fpm of 1 and to average out just above 1 (Van Heuven, Mandera, Keuleers, & Brysbaert, 2014) (Fig. 2.1 with caption from Van Heuven et al. (2014)).

2.1.2.4.4 Plausibility

As plausibility can also have an effect on acceptability judgements above and beyond syntactic structure itself (Garnsey, Pearlmutter, Myers, and Lotocky (1997); Garnsey, Tanenhaus, and Chapman (1989); Tanenhaus, Carlson, and Trueswell (1989), though cf. Sprouse (2008)), stimuli were also checked for this factor. This was done by running a separate web-based plausibility judgement task using Qualtrics®, where ten native Greek speakers - none of whom took part in the main study - were asked to rate the stimuli presented on a 5-point plausibility scale, with 1 being completely implausible or impossible and 5 be-

[8]Word frequency list based on a corpus of approximately 6,100 unique Modern Greek subtitle files from popular films and television series. Given the dominance of American cinema worldwide, the vast majority of these subtitles originate from USA productions (71.7%), with the remaining third encompassing UK (9.5%) and non-English productions, mainly German, French and Spanish films and television series (18.8%). The corpus is, thus, based on translation transcripts of the English language and not Modern Greek per se. This is, however, common practice with frequency corpora in various languages worldwide (e.g. Dutch: Keuleers, Brysbaert, and New (2010); Chinese: Cai and Brysbaert (2010); etc.) and is considered the most accurate estimate to date. The version used here is the so called 'SUBTLEX-GR_restricted', where the original raw material was subjected to cross-checking against a Modern Greek spell-checker in order to exclude optical character recognition spelling errors, resulting in a 'cleaner' corpus (Dimitropoulou, Duñabeitia, Avilés, Corral, & Carreiras, 2010).

Figure 2.1: Frequency effect on reaction times

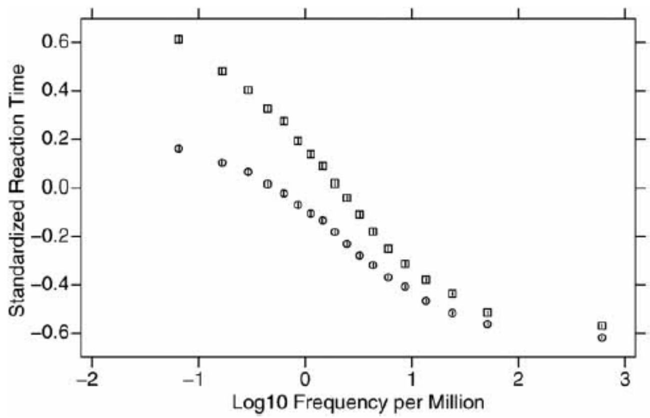


Figure 1. *The course of the word frequency effect in mean standarized reaction times from the British Lexicon Project (squares) and the English Lexicon Project (circles). The standard errors are represented by whiskers.*

ing completely plausible or realistic. A grammatical representative of each stimulus group (case-matched P-pied-piped sluice) and all grammatical fillers (64 in total) were assessed in this manner. In addition, an equal number of completely acceptable and plausible stimuli, as well as completely acceptable but implausible stimuli (altogether 128) were also included in the study as controls and so as to not bias participants towards one or the other end of the plausibility rating scale. In order to use more widely accepted plausible and implausible stimuli, the additional stimuli used were translated from a study by Ferreira (2003) in which they were normed by 100 participants. These were further embellished for additional complexity and sentence length to match the original experimental and filler items. Following Ferreira (2003), the 64 implausible sentences were divided into 32 reversible, semantically biased sentences (e.g. *The man bit the dog.*) and 32 non-reversible semantically implausible sentences (e.g. *The apron wore the chef.*). As the original study contained 24 of each, an extra 8 items were created based on the Ferreira (2003) sentences. For a full list, see Appendix (B.1.1.3).

The results showed all experimental stimuli to be judged as plausible (Mean: 4.80; SD: 0.21;) and not significantly different from either the grammatical fillers (Mean: 4.78; SD: 0.15; $p > .05$, *n.s.*) or plausible control stimuli (Mean: 4.88; SD: 0.2, $p > .05$, *n.s.*). Furthermore, they were significantly more plausible than the implausible control stimuli (Reversible Implausible Stimuli: Mean: 1.80; SD: 0.82; $p < .001$; Non-reversible Implausible

Stimuli: Mean: 1.24; SD: 0.7; $p < .001$).

2.1.2.5 Fillers

In addition to the experimental stimuli, 128 filler items were used. These were equally divided between simple active, simple passive, active embedded questions, passive embedded questions, VP-Ellipsis and DP-Ellipsis sentences. Half were ungrammatical and half grammatical.

2.1.2.6 Comprehension Questions

In order to make sure that participants paid adequate attention to the task at hand, comprehension questions were included for 2/3 of the experimental items and half of the grammatical fillers. This translated to 42 questions for the experimental items and 32 questions for the fillers. Half were false, half true. Regarding the experimental items, half of the questions concerned the matrix clause and half the embedded clause.

2.1.3 Results

Prior to analysis, the data were cleaned following the exclusion criteria detailed in section 2.1.2.3, leaving the data of 84 subjects. All responses made under less than $100ms$ (1 in total) were also excluded as erroneous. Outliers were calculated per condition per participant at $2.5 * SDs$ from the Mean and Windsorised (i.e. replaced by the minimum or maximum allowable value)[9]. In this fashion, $< 1\%$ of the total data points were replaced. After practice items and fillers were excluded, this left a total of 5,376 data points or 672 data points per experimental condition.

A maximal linear mixed effects (henceforth LME) model of analysis following Barr, Levy, Scheepers, and Tily (2013) was applied to the data using R's lme4 package with the default optimizer; the three factors of Case-matching (±), P-stranding (±) and Sluicing (±) served as fixed effects with subjects and items as random effects (with random slopes and intercepts assumed for each)[10]. The data was not normally distributed (Shapiro-Wilk all p's $< 8.9e-$

[9]There appears to be divided opinion in the scientific community over Windsorisation vs. deletion or 'trimming' of outlying data points, with both improving upon the rate of false positives or false negatives, i.e. making the analysis more robust compared to no data adjustment (Chen & Dixon, 1972). The two approaches do not appear to be vastly different in their output, but rather their difference is more philosophical, i.e. the logic of Windsorisation is that the researcher acknowledges these more extreme data points as potentially representing real data trends, particularly if they occur more than once for a particular item. By replacing them with the most extreme allowable value, they are acknowledged, but curbed so that they do not excessively skew the dataset. Deletion, on the other hand, simply ignores all outliers, including what could potentially be representative values. One clear advantage of Windsorisation over deletion is practical in nature, specifically that it does not add missing values to the dataset as deletion would and which can wreak havoc on analyses.

[10]This allows the model to take into account variance among subjects and items, as well as how each factor separately and in interaction may have differed per participant and/or item.

11), however general linear multivariate models are considered robust against violations of normality with skewed or kurtotic distributions having minimal effect on results (Donaldson, 1968; Glass, Peckham, & Sanders, 1972). However, it should be noted that different LME models were also run on Log10 transformations, z-scores and Log10-transformed z-scores of raw results, whilst also using different optimizers (*optimx*), each of which may have given better results had data distribution been an issue, however there was no significant difference between model fit results. As such, results of only the model with raw answers as dependent variable are presented here in text as they are the easiest to follow, particularly in barplots, with the rest shown in Table 2.7.

Mean results per condition and SEM are shown in Figure 2.2 and Table 2.6. The model showed significant two- and three-way interactions, as well as main effects for all three fixed factors, presented in more detail below, with post-hoc comparisons (Tukey-adjusted) for significant two- and three-way interactions carried out by comparing estimated marginal means (EMMs) via the "emmeans" R package. The results of the LME model showed a main effect for the fixed effects factor of a) Case-matching ($t > 17.67; p < 8.88e-16$[11]) with case-mismatching conditions being judged overall as significantly less acceptable than case-matching ones (Mean difference = 2.19); b) P-Stranding ($t > 14.74; p < 1.1e-16$), with P-pied-piping conditions being overall significantly more acceptable than P-stranding ones (Mean difference = 1.38); and c) Sluicing, ($t > 12.13; p < 6.88e-15$), with sluicing conditions being overall more acceptable than their overt counterparts (Mean difference = 1.04). These main effects were modulated by a significant two-way interaction between the factors of Case-matching and Sluicing ($t > 4.98; p < 6.5e-07$), with post-hoc comparisons showing that - when P-Stranding was collapsed across - a) the difference between case-matching and case-mismatching conditions was significant both overtly (Mean difference: $1.85; p < .0001$) and under sluicing (Mean difference: $2.60; p < .0001$), however the size of this difference was significantly smaller overtly than under sluicing (Mean size difference: $.75; p < .0001$); and b) the difference between overt and sluicing conditions was significant both for case-matching (Mean difference: $1.42; p < .0001$) and case-mismatching (Mean difference: $.67, p < .0001$), with both becoming significantly better under sluicing. The model additionally showed a significant two-way interaction between P-Stranding and Sluicing ($t > 10.48; p < 4.44e-16$), with post-hoc comparisons showing that, when Case-matching is averaged over, the difference between P-pied-piping and P-stranding was significant both overtly (Mean difference: $2.40; p < .0001$) and under sluicing (Mean difference: $0.48p < .003$), however the size of this difference was significantly larger overtly than under sluicing (Mean size difference: $1.83; p < .0001$); furthermore, there was a significant difference between overt and sluicing conditions when paired with P-stranding (Mean difference: 1.32; $p < .0001$), but not when paired with P-pied-piping (Mean difference: $0.01; p > .856, n.s.$). Together, these

[11] All p-values reported henceforth were calculated with Satterthwaite-approximated degrees of freedom.

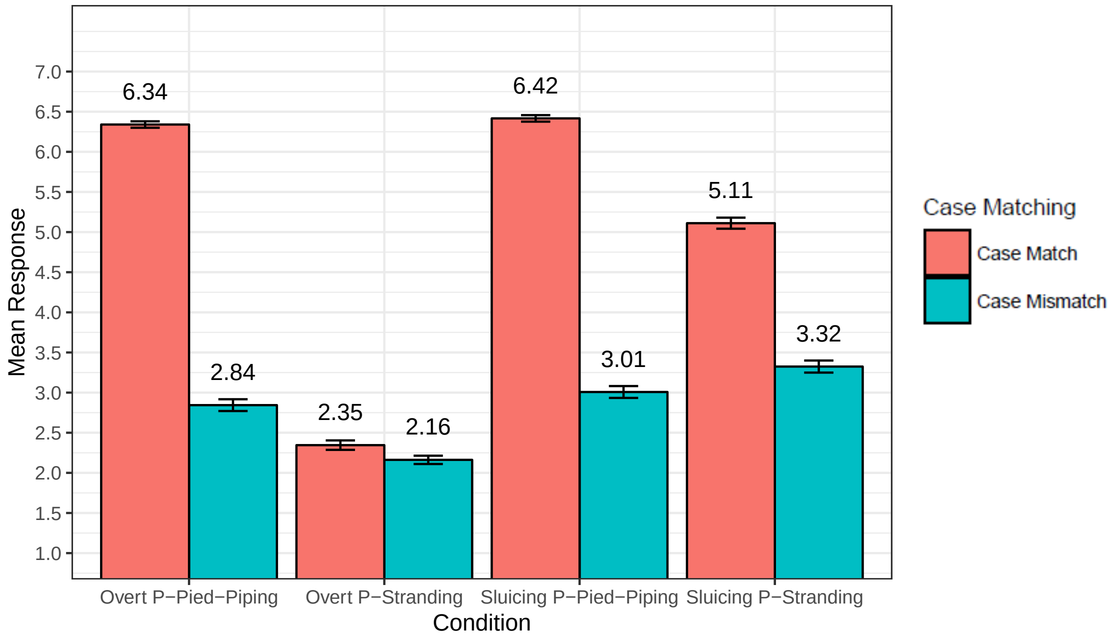
results indicate that - when Case-matching is not taken into account - although P-pied-piping is still significantly better than P-stranding both overtly and under sluicing, sluicing makes P-stranding significantly better than it is overtly, with no change in acceptability results for P-pied-piping. The final significant two-way interaction found was between the factors of P-stranding and Case-matching ($t > 13.18; p < 1.85e-16$), with post-hoc comparisons showing that when the factor of Sluicing is held constant, although the difference between P-pied-piping and P-stranding is significant for case-matching conditions (Mean difference: $2.65; p < .0001$), it is not significant for case-mismatching ones (Mean difference: $0.17; p > .484, n.s.$). With respect to the difference between case-matching and case-mismatching conditions, this difference remained significant for both P-pied-piping (Mean difference: $3.45; p < .0001$) and P-stranding conditions (Mean difference: $0.98; p < .001$) with an overall significantly larger difference for P-pied-piping conditions than P-stranding ones (Mean size difference: $2.47; p < .0001$).

The LME model also found a significant three-way interaction between all three variables. This complex interaction was carefully teased apart with post-hoc comparisons, taking each level of one of the variables separately from the other and checking how the interaction of the remaining two variables plays out each time. Specifically, when we focus on each level of Case-matching separately, for case-matching conditions the difference between P-pied-piping and P-stranding was significant both overtly (Mean difference: $4.08; p < .00001$) and under sluicing (Mean difference: $1.48; p < .0001$), however the size of this difference was significantly greater overtly (the largest difference between any minimal pair comparison, in fact) than under sluicing (Mean size difference: $2.61; p < .0001$). On the other hand, when we focus purely on case-mismatching conditions, although the difference between P-pied-piping and P-stranding was again significant overtly (Mean difference: $0.66; p < .0001$), it was not significant under sluicing (Mean difference: $0.31; p > .107, n.s.$). The factor of Sluicing thus affected the size of this difference, with it being significantly greater overtly than under sluicing (Mean size difference: $0.36; p < .01$), but also the direction of this difference (Mean size difference when taking into consideration difference in direction: $0.97; p < .0001$), with case-mismatching with P-stranding being significantly *less* acceptable than with P-pied-piping overtly, but significantly *more* acceptable than with P-pied-piping under sluicing. Case-mismatching was in fact the least acceptable when paired with P-stranding overtly and the most acceptable when paired with P-stranding under sluicing. When we focus now on each level of P-Stranding separately, for P-pied-piping conditions there is a significant difference between case-matching and case-mismatching both overtly (Mean difference: $3.50; p < .0001$) and under sluicing (Mean difference: $3.60; p < .0001$), with almost no difference in the size of this difference overtly and under sluicing (Mean size difference: $0.09; p > .393, n.s.$). When investigating P-stranding conditions, on the other hand, although there is a very slight difference between case-matching and case-mismatching overtly, this difference is not significant (Mean difference: $0.13; p > .726, n.s.$). Under sluicing, however, the difference between case-matching and case-mismatching conditions is significant (Mean difference: $1.59; p < .0001$), with the factor of Sluicing thus significantly affecting the size of this difference between case-matching and case-mismatching conditions (Mean size difference: $1.46; p < .0001$). Together with the previous results, this indicates that although overtly P-stranding together with case-mismatching is slightly less acceptable than P-stranding with case-matching, this is not a significant difference, with the two ungrammaticality effects having a slight numerically additive effect on ratings, but not a significantly important one. Under sluicing, on the other hand, there does appear to be a significant additive effect of the two ungrammaticality effects on ratings, with P-stranding together with case-mismatching being significantly worse than P-stranding alone. Let us now examine the difference between overt and sluicing conditions for P-pied-piping and P-stranding. Specifically, when we focus on P-pied-piping conditions only, there is no significant difference between overt and sluicing conditions when paired with either case-matching (Mean difference: $0.164; p > .496, n.s.$), nor case-mismatching (Mean difference: $0.07; p > .996, n.s.$). On the other hand, for P-stranding conditions there is a significant difference between overt and sluicing conditions both when paired with case-matching (Mean difference: $2.77; p < .0001$) and with case-mismatching (Mean difference: $1.305; p < .0001$), with the size of this difference being significantly greater with case-matching than case-mismatching (Mean size difference: $1.60; p < .0001$). Together with the significant two-way interaction we saw above between P-Stranding and Sluicing, this lends further evidence to the hypothesis that P-stranding becomes significantly better under sluicing than it is overtly, with no such amelioration difference for P-pied-piping. This also helps to further explain another significant two-way interaction seen above, namely that of Case-matching with Sluicing, indicating that the reason case-matching overall becomes significantly better under sluicing is predominantly due to P-stranding being significantly better with case-matching under sluicing, not P-pied-piping. Finally, and repeating some of the already examined differences for symmetry, when examining each level of Sluicing separately, for overt conditions, there was a significant difference between P-pied-piping and P-stranding conditions when paired with both case-matching (Mean difference: $4.08; p < .0001$) and case-mismatching (Mean difference: $0.71; p < .001$), with the size of this difference being significantly greater for case-matching conditions (Mean size difference: $3.37; p < .0001$). Under sluicing, on the other hand, although the difference between P-pied-piping and P-stranding conditions was significant when paired with case-matching (Mean difference: $1.48; p < .0001$), it was not significant when paired with case-mismatching (Mean difference: $0.53; p > .107, n.s.$).

For comparison purposes, Mean acceptability judgements are also presented for grammatical vs. ungrammatical fillers in Figure 2.3, with responses indicating that overall participants tended to avoid the absolute ends of the scale, even with what should be unequiv-

Figure 2.2: Experiment 1: Mean Acceptability Ratings Barplot

Mean response per condition with SEM error bars. In this and all following barplots Non-Sluicing conditions have been renamed 'Overt' to save space.



ocally acceptable and unacceptable sentences, lending more support to the idea that the results we found for overt P-stranding and case-mismatching conditions do indeed represent 'unacceptable' judgements, albeit with some gradation.

Table 2.6: Experiment 1: Acceptability Ratings Numerical Summary

	Condition	Response	SD	SEM	95% CI
1	Case Matching Non-Sluicing P-Pied-Piping	6.340	1.046	0.040	0.079
2	Case Matching Non-Sluicing P-Stranding	2.345	1.528	0.059	0.116
3	Case Matching Sluicing P-Pied-Piping	6.417	1.051	0.041	0.080
4	Case Matching Sluicing P-Stranding	5.111	1.783	0.069	0.135
5	Case Mismatching Non-Sluicing P-Pied-Piping	2.844	1.902	0.073	0.144
6	Case Mismatching Non-Sluicing P-Stranding	2.162	1.336	0.052	0.101
7	Case Mismatching Sluicing P-Pied-Piping	3.008	1.909	0.074	0.145
8	Case Mismatching Sluicing P-Stranding	3.324	1.945	0.075	0.148

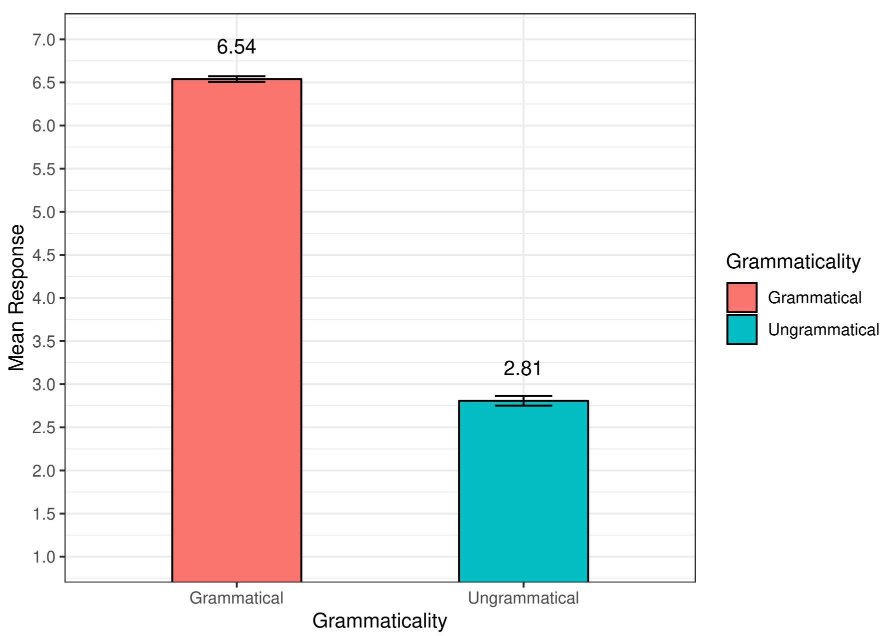
With respect to the overall model fit, based on the formulas derived by Nakagawa and Schielzeth (2013) and the R package 'piecewiseSEM', as created by Lefcheck (2016), we can

Table 2.7: Experiment 1: Main Effects & Interactions

		Raw Scores		z-Scores		$Log10$-Scores	
	Factor	t-value	p-value	t-value	p-value	t-value	p-value
1	Case-Matching	17.673	< 0.0001	24.670	< 0.0001	14.349	< 0.0001
2	P-Stranding	14.744	< 0.0001	16.459	< 0.0001	11.564	< 0.0001
3	Sluicing	12.134	< 0.0001	13.273	< 0.0001	11.532	< 0.0001
4	Case-Matching*P-Stranding	13.178	< 0.0001	15.001	< 0.0001	10.670	< 0.0001
5	Case-Matching*Sluicing	4.976	< 0.0001	5.775	< 0.0001	4.227	< 0.0001
6	P-Stranding*Sluicing	10.475	< 0.0001	12.765	< 0.0001	10.302	< 0.0001
7	Case-Matching*P-Stranding*Sluicing	4.781	< 0.0001	5.884	< 0.0001	4.855	< 0.0001

Figure 2.3: Experiment 1: Fillers Mean Acceptability Ratings Barplot

Mean response per condition with SEM error bars.



calculate the conditional[12] R^2 for the GLME model as shown in equation (2.1)[13].

$$R^2_{GLMM(c)} = \frac{\sigma_f^2 + \sum_{l=1}^{u} \sigma_l^2}{\sigma_f^2 + \sum_{l=1}^{u} \sigma_l^2 + \sigma_e^2 + \sigma_d^2} \tag{2.1}$$

This gives us a conditional R^2 of .66, meaning that this model, as presented, i.e. including both fixed and random effects factors, can account for approximately 66% of the variation in the data - a large effect size, following Sullivan and Feinn (2012) and adapted from Ferguson (2009). A power analysis, following Westfall, Kenny, and Judd (2014), estimated the study power at 1, meaning the same result would be obtained 100% of the time (i.e. a very robust result).

Finally, I would like to take the opportunity here to point out the results of a few further factor analyses which were secondarily included in the design, as mentioned in section 2.1.2.4, in order to test how these may have contributed to acceptability differences, based on previous theoretical explanations of PSG-defying languages. Specifically, with respect to the factor of Embedding Verb ([+Acc] vs. [-Acc.]), adding this to the LME model produced no significant main effect nor almost any interactions (all t's $< 1.22; p$'s$> .16, n.s.$); the exception was a significant two-way interaction of Case-matching with Embedding Verb ($t > 2.56; p < .011$), which post-hoc comparisons showed was due to the acceptability difference between [+Acc.] and [-Acc.] embedding verbs being significant for case-matching conditions (Mean difference: $0.31; p < .008$), but not case-mismatching ones (Mean difference: $0.09; p > .319, n.s.$). This interaction was further modulated by a significant three-way interaction of Case-matching, Sluicing and Embedding Verb ($t > 2.57; p < 0.010$), with post-hoc comparisons showing that this was due to the difference in acceptability between [+Acc.] and [-Acc.] embedding verbs being significant only for case-matching sluicing conditions, where [+Acc.] verbs were rated marginally significantly higher than [-Acc.] ones (Mean difference: $0.44; p < .055$), despite no such difference for case-matching conditions overtly, nor case-mismatching conditions overtly or under sluicing (all p's $> .813$). This would indicate that overall there does appear to be some minor acceptability advantage for sluicing with case-matching when the embedding verb also patterns with Accusative, however this does not extend to differences in P-stranding acceptability either overtly or under sluicing. Hence, although interesting as an observation, this is irrelevant to our story. With respect to the factor of Remnant Complexity (*which*-NP vs. bare *wh*-remnant), there was a significant main effect found ($t > 2.44, p < .015$), with bare *wh*-remnants (Mean: 4.00) being on average slightly more highly rated than *which*-NP remnants (Mean: 3.40)

[12]Where the conditional R^2 refers to the model fit including both fixed effects and random effects factors, as opposed to the marginal R^2 which would include only fixed effects.

[13]Where σ_f^2 is the fixed effects factor variance; σ_l^2 is the lth random factor variance; u is the number of random factors; σ_e^2 is the additive dispersion variance and σ_d^2 is the distribution-specific variance, which together make up the residual error variance, σ_ϵ^2, in GLMMs (Nakagawa & Schielzeth, 2010).

across all conditions, with this effect not being modulated by any significant interactions (all t's$< 1.60, p$'s$> .11, n.s.$). With respect to the factor of Number (singular vs. plural), there was no significant main effect nor interaction when it was entered into the LME model (all t's < 1.87; all p's $> 0.11, n.s.$). Finally, the factor of PP Type ('me' vs. 'se' vs. 'jia' vs. 'apo') showed no main effect nor significant interactions (all t's $< .8$; all p's $> .39, n.s.$) when added to the model.

Before these results are discussed any further, however, it should be noted that there were a number of participants who consistently rated P-stranding under sluicing as less than acceptable (i.e. below 4.0/7.0)). Although these were not many in number (N = 9), upon closer examination it became apparent that a common characteristic was that they originated from the North of Greece and more specifically the area in and around Thessaloniki (Greece's second largest city). Although not differing in their ratings of other conditions, these participants rated P-stranding under sluicing at 3.7 on average (Mean: 3.68; SD: .29), as opposed to participants from other areas who rated it at 5.2 on average (Mean: 5.15; SD: .13). Given how unequal the two sample sizes were between people originating from Thessaloniki and the rest of Greece, it was impossible to accurately calculate whether region of origin had a significant effect on acceptability rating. Numerically, however, the two appear substantially different.

2.1.4 Discussion

The results of this study show that P-stranding under sluicing in Greek is evidently significantly more acceptable than it is overtly. However, it is also not as acceptable as P-pied-piping, either under sluicing or overtly. These results are not predicted by any theoretical approach to sluicing. With respect to the pointer account as argued for by Nykiel (2013), the difference between P-stranding and P-pied-piping under sluicing is also unexpected, particularly given the stimuli contained both complex and simplex remnants and correlates, with no difference found between them.

As expected, all of the case-mismatching conditions were significantly less acceptable than their case-matching counterparts, whether overtly or under sluicing, and whether coupled with P-pied-piping or P-stranding, something predicted by all theories. Interestingly, however, there was a significant interaction of case-matching with sluicing, with case-mismatching being significantly more acceptable under sluicing. This difference, along with the fact that it is not as profound an amelioration as for P-stranding, is perplexing for any of the theories we have thus far described. The same is true with respect to the overt difference between case mismatching with P-pied-piping vs. P-stranding, as all theories would predict the two should be equally unacceptable and at floor level, whereas in fact it would appear as though the two violations have a slight additive effect on acceptability deterioration compared to each violation separately.

Furthermore, there appeared to be no contributing effect to these results of any of the alternative factors which have been - or could be - argued to affect the acceptability of P-stranding under sluicing. With respect to the embedding verb, there was no significant effect of this factor when added to the model, indicating that any apparent acceptability of P-stranded remnants cannot be attributed to a grammatical illusion of some sort. It should be noted that there was a significant three-way interaction of embedding verb with case-matching and sluicing, such that the difference between case-matching and case-mismatching under sluicing was larger for [+Acc.] verbs compared to [-Acc.]-patterning verbs, indicating that there may be a temporary grammatical illusion effect occurring, but only with respect to the factor of case-matching under sluicing and not P-stranding. With respect to the form of the remnant, although there was a significant main effect for this factor, upon closer examination it seems driven by bare *wh*-remnants being on average slightly *more* acceptable than *which*-NP-remnants across conditions (i.e. the opposite direction of what the existing literature would predict), however it is possible this effect is circumstantial. Importantly, however, there was no significant interaction between this factor and any others. Finally, with respect to the number of the remnant/correlate, there was no significant difference found between singular and plural. If the amelioration of case mismatching under sluicing found in both this and the previous (MSc) experiment was potentially due to participants misreading the final letter of the stimuli as the case-matching Accusative suffix (*-n*) vs. the case-mismatching Nominative suffix (*-s*), then this would be less likely to occur for a) plural number and b) *which*-NPs, as in both cases there are more than 1 orthographic string differences compared to bare singular *wh*-remnants. Neither of these factors, however, appeared to affect the acceptability of case-mismatching.

One last, interesting side observation here is that despite unnecessary material being repeated in the overt continuation conditions, there was no difference whatsoever between Case-matching P-pied-piping with an overt continuation vs Case-matching P-pied-piping under sluicing. This is somewhat surprising in light of the 'repeated name penalty' which has been argued to affect processing and thereby acceptability in the literature (Almor, 1999; Garrod et al., 1994; Gordon et al., 1993).

2.2 Experiment 2: P-Stranding out of Islands under Sluicing

2.2.1 Experimental Outline, Aim and Predictions

Based on the results of the previous experiment we can conclude that, firstly, Greek does indeed appear to be a non-P-stranding language overtly, but also, more importantly, that P-stranding does also appear to be significantly more acceptable under sluicing compared to overtly, albeit not as much as P-pied-piping. These results are not clearly in line with any of the previously mentioned models, either processing- or theoretical syntax-based.

One possibility that deserves to be examined here in more detail is the idea that PPs may in fact be PF-islands in Greek, similarly to how Almeida and Yoshida (2007), Leung (2014) and - to some extent - Sato (2011) argued that this was the reason P-stranding was supposedly allowed under sluicing for Brazilian Portuguese, Emirati Arabic and Bahasa Indonesian, respectively. Of course, as mentioned earlier, this kind of logic does beg the question of a) why this would be true for Greek and these other languages, but not all languages that disallow P-stranding overtly, and b) why P-stranding would not also be allowed in other cases of ellipsis targeting PP-containing constituents, such as VP-ellipsis. Leaving these issues aside for now, however, let us simply assume that PPs may indeed be some form of island ameliorated by sluicing. If this is true, then one would expect to see P-stranding amelioration behave similarly to island amelioration under sluicing.

Interestingly, however, we do not have any such data against which to compare P-stranding amelioration. That is to say, even though sluicing has been widely accepted since its naissance to ameliorate a number of islands by default, no study has thus far examined whether they are indeed entirely acceptable under sluicing, as theory would predict, or whether they are somewhat ameliorated compared to overt extraction, but not to the same degree of acceptability perhaps as regular sluice remnants. The latter case, in particular, would presumably result in a situation reminiscent of what we observed for P-stranding under sluicing in Experiment 1 as compared to P-pied-piping; if this is true, then this might strengthen an argument for PPs, for some reason, behaving like PF-islands for some languages.

In order to examine both these questions, i.e. establish to what degree islands are indeed ameliorated under sluicing, as well as how this compares to the P-stranding amelioration observed in the previous study, in the following experiment I once again crossed the factors of Sluicing (±, i.e. Overt continuation vs. Sluicing) and P-Stranding (±, i.e. P-pied-piping vs. P-stranding), but this time in the context of islands, resulting in the Conditions presented in Table 2.8. In both this and the following study, I dropped the factor of Case-matching, focusing instead only on P-Stranding, which is the main subject of interest of this thesis. Practically, this meant that the conditions in Experiment 2 were very similar to those of Experiment 1, however here they all involved extraction out of an island, as in (58). By comparing the results of the two experiments, this design allows us to examine one further aspect of sluicing which has not been previously considered, and that is the interaction of P-stranding and islandhood. One aspect of this interaction concerns sluicing amelioration, more specifically whether P-stranding and/or P-pied-piping out of a supposed island are ameliorated to a different degree to each other and compared to when there is no island; the other aspect of this interaction concerns P-stranding overtly and whether it is worse when extracting out of an island compared to when there is no island.

(58) To oti i Maria chorepse me kapjon ine
*Det.*N.NOM *that Det.*F.NOM *Maria.*F.NOM *danced.*3SG*.* *with someone.*M.ACC *is.*3SG
gnosto, alla de thimame me pjon akrivos ine
*well-known.*N.NOM *but* NEG *remember.*1SG *with who.*M.ACC *exactly is.*3SG
gnosto oti chorepse.
*well-known.*N.NOM *that danced.*3SG*.*
'That Maria danced with someone is well-known, but I cannot remember with who exactly it is well-known that she danced.'

Having outlined the motivations for this experiment, let us examine the predictions that each theoretical approach we have seen thus far would make for this study's design. Regardless of the approach one takes here, the one common thing which all theories would expect is that movement of a full PP out of an island overtly should be considered unacceptable. However, with respect to the other three conditions, as well as how they compare to those of the previous experiment, each approach would likely make slightly different predictions.

Let us start with a PF-Deletion Hypothesis, with the reader being referred to Table 2.9 to make these predictions clearer. First of all, if islands are considered to be ameliorated under sluicing, for instance due to an ungrammatical-feature-bearing island node being deleted, then one would expect them to be as acceptable as regular sluices. In other words, in this study, on the one hand, the P-pied-piping condition should be significantly better under sluicing than overtly; and in comparison with the previous study, on the other hand, both studies should show similar acceptability of this P-pied-piping under sluicing condition. With respect to P-stranding, Merchant (2001; and subsequently) proposes that the PSG should hold true even when the sluice is considered to contain an island (Merchant, 2001, pp. 105-107), with many such examples being provided for Greek amongst other languages. That is to say, regardless of the complexity or grammaticality of the structure it is being extracted out of, a P-stranding remnant should still be considered unacceptable. Firstly, therefore, this statement entails that P-stranding should always be unacceptable and at floor level, regardless of Sluicing or Islandhood. Together with the prediction on P-pied-piping, this translates to a predicted significant interaction between the factors of P-Stranding and Sluicing, such that P-pied-piping is unacceptable overtly, but ameliorated under sluicing, and P-stranding is simply always unacceptable. Secondly, with respect to the interaction of Sluicing and Islandhood, this statement would entail there should be no clear difference between P-stranding under sluicing in the context of islands vs. in the absence of islands, with both being completely unacceptable and at floor-level. Given the theory does not predict different levels or a gradient of acceptability anywhere nor any form of additive unacceptability, it would presumably predict that P-stranding overtly should also always be

Table 2.8: Experiment 2 Conditions and Examples

(59) To oti i neari mathitria krivotan apo
Det.N.NOM *that Det*.F.NOM *young*.F.NOM *student*.F.NOM *was.hiding*.3SG *from*
kapjous sti scholiki avli itan gnosto, ala
someone.M.ACC *in.Det*.F.ACC *school*.F.ACC *yard*.F.ACC *was*.3SG *well-known*.N.NOM *but*
kanis de borouse na thimithi ...
no.one.M.NOM NEG *could*.3SG SUBJ *remember*.3SG
'That the young student was hiding from some people in the school yard was well-known, but no one could remember ...'

	Condition	Example
1	Non-Sluicing, P-Pied-Piping	... apo pjous itan gnosto oti i *from who*.M.ACC *was*.3SG *well-known*.N.NOM *that Det*.F.NOM neari mathitria krivotan. *young*.F.NOM *student*.F.NOM *was.hiding*.3SG '*...from who it was well-known that the young student was hiding.'
2	Non-Sluicing, P-Stranding	... pjous itan gnosto oti i *who*.M.ACC *was*.3SG *well-known*.N.NOM *that Det*.F.NOM neari mathitria krivotan apo. *young*.F.NOM *student*.F.NOM *was.hiding*.3SG *from* '*...who it was well-known that the young student was hiding from.'
3	Sluicing, P-Pied-Piping	... apo pjous *from who*.M.ACC '...from who.'
4	Sluicing, P-Stranding	... pjous *who*.M.ACC '...who.'

at floor level, whether in the context of islands or not. As such, all overt conditions containing a) P-stranding out of an island; b) P-pied-piping out of an island; and c) P-stranding with no island should all be equally unacceptable and at floor level.

Table 2.9: PF-Deletion Acceptability Predictions for P-Stranding, Sluicing & Islands

		Island	No Island
P-Pied-Piping	Overt Continuation	×	✓
	Sluicing	✓	✓
P-Stranding	Overt Continuation	×	×
	Sluicing	×	×

Notes: The crosses and checkmarks depict floor- and ceiling-level acceptability, respectively. The first, 'Island', column refers to the predictions for this experiment, and the second, 'No Island', column to those of the previous experiment for comparison.

Now, building on top of this original PF-Deletion account, if we make the additional hypothesis, along the lines of Almeida and Yoshida (2007), Leung (2014) and Sato (2011), that PPs are PF-islands in Greek, for whatever reason, then the table above becomes that in 2.10, with P-stranding now also being ameliorated under sluicing, regardless of whether it appears in the context of an additional island or not. To explain this, recall that island effects have been argued to be due to the crossing of an island node bestowing an ungrammatical feature on this node, which, when pronounced, triggers a derivational crash. Thus, if we have P-stranding out of an island, we would presumably have two ungrammatical features, one associated with extraction out of the PP-island, and one with extraction out of the other island. If neither of these ungrammatical island features are pronounced thanks to sluicing, then we would anticipate sluicing to fully ameliorate such a structure across the board. This would translate to a main effect of the factor of Sluicing, such that sluicing conditions are significantly better overall compared to overt conditions, with no significant interaction between factors. To frame this also in terms of comparison with the previous experiment, we would anticipate that P-pied-piping with or without an island as well as P-stranding with or without an additional island, should all be identically acceptable under sluicing and at ceiling-level. This prediction of identical acceptability, as mentioned above, is due to the PF-Deletion Hypothesis not making any gradient acceptability predictions, simply stating that a structure is either acceptable or not.

Moving on to an LF-Copying account, given the lack of structure it posits at the *e*-site, this account naturally predicts island amelioration under sluicing, as mentioned in Chapter 1. Furthermore, its additional structural isomorphism requirement between correlate and

Table 2.10: PF-Deletion Acceptability Predictions for P-Stranding, Sluicing & Islands, with PPs as PF-Islands

		Island	No Island
P-Pied-Piping	Overt Continuation	×	✓
	Sluicing	✓	✓
P-Stranding	Overt Continuation	×	×
	Sluicing	✓	✓

Notes: The crosses and checkmarks depict floor- and ceiling-level acceptability, respectively. The first, 'Island', column refers to the predictions for this experiment, and the second, 'No Island', column to those of the previous experiment for comparison.

remnant is technically restricted only to case-identity, not phrasal category-identity. Thus, unless we extend this requirement to phrasal category-identity, which would overgenerate as mentioned previously, P-stranding out of an island is also predicted to be acceptable under sluicing, similarly to P-pied-piping out of islands. The acceptability predictions of an LF-Copying account for this experiment and in comparison with the previous experiment would, therefore, be those depicted in Table 2.11, and would technically be identical to those of a PF-Deletion Hypothesis with PPs as PF-Islands. To reiterate, P-pied-piping, whether in the context or not of islands, as well as P-stranding, again in the context or not of islands, should be identically acceptable under sluicing and at ceiling-level. In other words, for this study we would predict a significant main effect of Sluicing and no significant interaction between the factors of Sluicing and P-Stranding.

Table 2.11: LF-Copying Acceptability Predictions for P-Stranding, Sluicing & Islands

		Island	No Island
P-Pied-Piping	Overt Continuation	×	✓
	Sluicing	✓	✓
P-Stranding	Overt Continuation	×	×
	Sluicing	✓	✓

Notes: The crosses and checkmarks depict floor- and ceiling-level acceptability, respectively. The first, 'Island', column refers to the predictions for this experiment, and the second, 'No Island', column to those of the previous experiment for comparison.

Moving on to a pointer mechanism account with no structure posited at the *e*-site, one would predict that P-stranding, P-pied-piping and P-pied-piping out of islands should be uniformly acceptable under sluicing, thanks to this lack of structure. With respect to the pointer account of Nykiel (2013) more specifically, things are a little more complicated, as can be seen in Table 2.12. As mentioned in Chapter 1 (1.3.3), Nykiel specifically proposes an approach following Ariel's Accessibility Theory, such that complex correlates are more easily retrieved via less complex remnants, and simplex correlates via more complex remnants. Complex correlates, Nykiel argues, such as *some minister*, are always referred to via *which*-NP remnants and not bare *wh*-XPs, with the latter only being able to refer to simplex correlates, such as *someone*. [14]. As such, given they refer back to a richer PP correlate in terms of information, *which*-NP remnants are able to appear without their P. Since this would make the remnant less complex and thereby more well-suited to referring to a complex correlate, I argued in Chapter 1 that this should translate to P-stranding *which*-NP remnants being *more* acceptable than P-pied-piping *which*-NP remnants, which is the opposite of what Nykiel's studies showed. Conversely, one would expect that *wh*-XP remnants would allow P-stranding to a much smaller degree, with P-pied-piping being more well-suited to referring to the simpler PP correlate. As Nykiel argues for a gradient in acceptability based on processing costs, we might thus predict the following gradient: sluicing with P-stranding *which*-NPs > sluicing with P-pied-piping *which*-NPs > sluicing with P-pied-piping *wh*-XPs > sluicing with P-stranding *wh*-XPs, with the advantage of *which*-NPs over *wh*-XPs arising from more shared features between correlate and remnant in the former (following Nairne (2006)). To check whether such a gradient is indeed present in Greek, as in Experiment 1, here half of the stimuli had a *which*-NP remnant with a complex correlate and half a *wh*-XP remnant with a simplex correlate. If the above is true, we would anticipate a significant interaction between the factors of P-Stranding and Complexity of Remnant here. In terms of the two overarching factors being crossed here, i.e. Sluicing and P-Stranding, one would anticipate that both P-pied-piping and P-stranding should be overall acceptable under sluicing and unacceptable overtly, i.e. we would anticipate a main effect of Sluicing. Given this is a processing-based account with respect only to how sluicing itself is handled, with the theory not explicitly dealing with overt constructions, it would thus be reasonable to assume that a pointer account would make the same predictions as an LF-copying or PF-Deletion account with respect to how P-stranding and islands should be dealt with in the context of non-sluicing continuations, i.e. when combined, P-stranding out of an island should not have a compounding effect, instead being judged at floor level and equally to both overt P-pied-piping out of an island and overt P-stranding with no island. As such, when compared with the results of the previous study, the theory should predict a) a significant two-way interaction between P-Stranding and Sluicing, which should be modulated

[14] As mentioned in Chapter 1, it is unclear here why this is considered to be in accordance with Ariel's Accessibility Theory, given that theory would predict the opposite pairing, with complex correlates being referred back to by simplex remnants and vice versa.

by b) a significant three-way interaction between P-Stranding, Sluicing and Complexity of Remnant, but *not* by c) a significant three-way interaction between P-Stranding, Sluicing and Islandhood.

Table 2.12: Pointer Account (Nykiel, 2013) Acceptability Predictions for P-Stranding, Sluicing & Islands

		Island		No Island	
		Which-NP	*Wh*-XP	*Which*-NP	*Wh*-XP
P-Pied-Piping	Overt Continuation	✕	✕	✓	✓
	Sluicing	✓	✓	✓	✓
P-Stranding	Overt Continuation	✕	✕	✕	✕
	Sluicing	✓	?/✕	✓	?/✕

Notes: The crosses and checkmarks depict floor- and ceiling-level acceptability, respectively. The first set of columns, marked 'Island', refer to the predictions for this experiment, and the second set of columns, marked 'No Island', to those of the previous experiment for comparison.

For a summary of these predicted effects by each account, please see Table 2.13 below.

Table 2.13: Experiment 2 Significance Predictions

	Prediction of Significance	PF-Deletion	PF-Deletion (PPs as Islands)	LF-Copying	Pointer
1	Sluicing	✕	✓	✓	✓
2	P-Stranding	✓	✕	✕	?/✓
3	Sluicing*P-Stranding	✓	✕	✕	?/✕
4	Sluicing*P-Stranding*Remnant-Type	✕	✕	✕	✓

Notes: As in Experiment 1, 'Remnant Type' is an additional factor controlled for here, as will be made clear in the next section.

2.2.2 Methods

As with Experiment 1, a web-based acceptability judgement task was conducted via Ibex Farm®, using the same presentation method, instructions and rating scale. In the interest of brevity only those aspects of the experiment differing from Experiment 1 will be detailed here. The factors of Sluicing (±) and P-Stranding (±) were crossed once more, with the addition that each condition also involved overt or covert extraction out of an island, resulting in the conditions presented in Table 2.8 above, repeated below in Table 2.14 for ease. Out of the 85 native speakers recruited, using the same exclusion criteria as in Experiment 1, the data of only 60 were used.

2.2.3 Materials

As the study contained fewer conditions, fewer stimuli groups - and consequently fillers - were presented than in Experiment 1. There were 32 stimuli groups and 64 fillers in total, resulting in 8 data points per condition per participant, as in Experiment 1. Where feasible, the same lexical items were used as in the first experiment, and where not, the new items were selected based on the same criteria as those in the previous experiment, i.e. following the same strict set of rules to control for confounding factors and to check for the effect of other factors previously proposed to explain P-stranding under sluicing. The main addition to these stimuli was that all contained supposed islands in the sluice region, rendering both overt conditions theoretically unacceptable. Half the stimuli contained relative clause islands (60) and the other half subject islands (61).

(60) Sto filladio odijion egrafe pos to neo
In.Det.N.ACC *leaflet*.N.ACC *instructions*.F.GEN *wrote*.3SG *that Det*.N.NOM *new*.N.NOM
farmako endiknite jia atoma pou paschoun apo kapjon
drug.N.NOM *is.meant*.3SG *for people*.N.ACC *that suffer*.3SG *from some*.M.ACC
mikita, alla den itan ksekatharo (apo) pjon mikita.
fungus.M.ACC *but* NEG *was*.3SG *clear*.N.NOM *(from) which*.M.ACC *fungus*.M.ACC
'In the instructions, it said that the new medicine was meant for people who were suffering from a specific fungus, but it was not clear (from) which fungus.'

(61) To oti i dikijoros erjazotan jia kapjous
Det.N.NOM *that Det*.F.NOM *lawyer*.F.NOM *was.working*.3SG *for someone*.M.ACC
sti mafia itan veveos gnosto, alla kanis den
in.Det.F.ACC *mob*.F.ACC *was*.3SG *of.course known*.N.NOM *but no.one*.M.NOM NEG
tolmouse na ksestomisi (jia) pjous akrivos.
dared.3SG SUBJ *say.out.loud for who*.M.ACC *exactly*

Table 2.14: Experiment 2 Conditions and Examples (Repeated)

(59) To oti i neari mathitria krivotan apo
Det.N.NOM *that* *Det*.F.NOM *young*.F.NOM *student*.F.NOM *was.hiding*.3SG *from*
kapjous sti scholiki avli itan gnosto, ala
someone.M.ACC *in.Det*.F.ACC *school*.F.ACC *yard*.F.ACC *was*.3SG *well-known*.N.NOM *but*
kanis de borouse na thimithi ...
no.one.M.NOM NEG *could*.3SG SUBJ *remember*.3SG
'That the young student was hiding from some people in the school yard was well-known, but no one could remember ...'

	Condition	Example
1	Non-Sluicing, P-Pied-Piping	... apo pjous itan gnosto oti i *from* *who*.M.ACC *was*.3SG *well-known*.N.NOM *that* *Det*.F.NOM neari mathitria krivotan. *young*.F.NOM *student*.F.NOM *was.hiding*.3SG '*... from who it was well-known that the young student was hiding.'
2	Non-Sluicing, P-Stranding	... pjous itan gnosto oti i *who*.M.ACC *was*.3SG *well-known*.N.NOM *that* *Det*.F.NOM neari mathitria krivotan apo. *young*.F.NOM *student*.F.NOM *was.hiding*.3SG *from* '*... who it was well-known that the young student was hiding from.'
3	Sluicing, P-Pied-Piping	... apo pjous *from* *who*.M.ACC '... from who.'
4	Sluicing, P-Stranding	... pjous *who*.M.ACC '... who.'

'That the lawyer was working for some people in the mob was of course well-known, but no one would dare say who exactly.'

Comprehension questions were created and presented as in Experiment 1, i.e. with half being false and half true, targeting either the main clause or the sluice, and presented after 2/3 of P-pied-piping sluices and P-stranding sluices. Fillers were kept the same between the two studies in order to have a common level of reference, and only half used for Experiment 2 given the smaller number of experimental items. Half of these were grammatical and half ungrammatical.

2.2.4 Results

Prior to analysis, the data was cleaned following the same principles as Experiment 1, i.e. excluding the data of participants who responded with less than 80% accuracy to all comprehension questions, wrote an incomprehensible paragraph in the introductory form and/or rated all sentences the same or inverted the rating scale. We were thus left with the data of 60 subjects. No responses were made below $100ms$. Outliers calculated per condition at $2.5 * SDs$ from the Mean were Windsorised (i.e. replaced by the minimum or maximum allowable value). A total of $< 2\%$ of the total number of data points were replaced in this fashion. After practice items and fillers were excluded, this left a total of 1,903 data points or approximately 478 data points per condition.

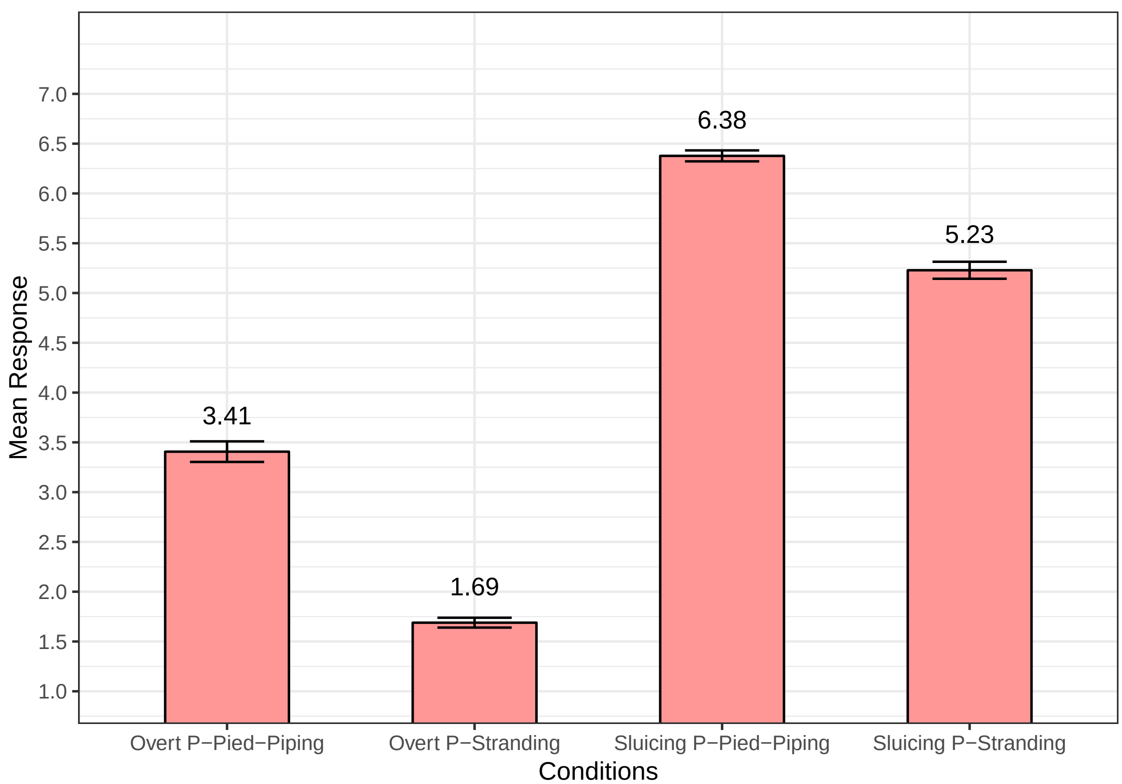
A maximal LME model with the fixed effects of sluicing (±) and P-stranding (±) and random effects of items and subjects (random slopes and intercepts assumed) (Barr et al., 2013) and default optimizer was fitted to the data. As in Experiment 1, the data was not normally distributed, and for this reason different LME models were run with raw responses, z-scores and $Log10$-transformed responses, but also using a different, stronger optimizer (*optimx*), each of which may have given better results had normality violation been a significant issue. As the full model including random slopes and intercepts for subjects and items failed to converge for the raw scores, as well as z-scores, the model output with $Log10$-transformed scores as dependent variable is presented here. These results and along with the model output for raw scores and z-scores as DV with only random intercepts are shown in Table 2.16.

Table 2.15 presents a numerical summary of raw scores and Figure 2.4 presents raw score Acceptability Means with SEM per condition. Raw scores are depicted here as these are easier to follow and compare with the other experiments.

The LME model with $Log10$-transformed scores as dependent variable converged showing a significant main effect for both factors and significant two-way interaction which was explored further with post-hoc comparisons (Tukey-adjusted) carried out by comparing estimated marginal means (EMMs) via the "emmeans" R package. Specifically the model showed a significant main effect for the factor of Sluicing ($t > 13.85; p < 1e - 17$), with sluices being rated on average much higher than overt island structures (Mean difference: 3.25), as expected by all theories. There was also a significant main effect of P-Stranding ($t > 7.45; p < 9.46e - 14$), with P-stranding conditions being overall less acceptable compared to P-pied-piping (Mean difference: 1.42). However, both these effects were modulated by a significant two-way interaction between the two factors of Sluicing and P-Stranding ($t > 2.51; p < .012$), with the difference between overt and sluicing conditions being significant both with P-stranding (Mean difference: $3.54; p < .001$) and P-pied-piping (Mean difference: $2.97; p < .001$), however the size of this difference was significantly greater for P-stranding than for P-pied-piping conditions (Mean size difference: $0.54; p < .001$). Hence, although both P-pied-piping and P-stranding were significantly more acceptable under sluicing compared to overtly thanks to island amelioration, the amelioration for P-stranding was significantly more noticeable.

Figure 2.4: Experiment 2: Mean Acceptability Ratings Barplot

Mean response per condition with SEM error bars.



Following the same process as in Experiment 1, we can calculate the model's fit following

Table 2.15: Experiment 2: Acceptability Ratings Numerical Summary

	Condition	Response	SD	SEM	95% CI
1	Non-Sluicing P-Pied-Piping	3.407	2.256	0.103	0.203
2	Non-Sluicing P-Stranding	1.689	1.075	0.050	0.097
3	Sluicing P-Pied-Piping	6.377	1.203	0.055	0.108
4	Sluicing P-Stranding	5.229	1.865	0.085	0.168

Table 2.16: Experiment 2: Main Effects & Interactions

		Raw Scores		z-Scores		$Log10$-Scores	
	Factor	t-value	p-value	t-value	p-value	t-value	p-value
1	P-Stranding	20.864	< 0.0001	21.739	< 0.0001	7.452	< 0.0001
2	Sluicing	43.373	< 0.0001	44.463	< 0.0001	13.852	< 0.0001
3	P-Stranding*Sluicing	4.263	< 0.0001	3.843	< 0.0001	2.514	< 0.0119

Nakagawa and Schielzeth (2013). This gives us a conditional R^2 of .76, meaning that this model, as presented, i.e. including both fixed and random effects factors, can account for approximately 76% of the variation in the data. A power analysis (following Westfall et al. (2014)) estimated the study power at 1, meaning the same result would be obtained 100% of the time (i.e. a very robust result).

Let us have a look now at the other factors included in the design which may have affected results. Out of the factors of embedding verb (±Acc.), remnant and correlate type (complex vs. simplex), remnant number (singular vs. plural) and P type (*apo* vs. *jia* vs. *me* vs. *se*), when each of these was added to the 2x2 LME model (keeping random slopes and intercepts assumed for the two fixed effects factors), none showed any main effects or significant two-way or three-way interactions (all t's$<$ 1.49; all p's$>$ .14, $n.s.$). As a side-note, numerically at least, there did appear to be a slight difference between RC- and subject-island extraction; overall subject-island extraction (Mean: 4.44) was rated more highly than RC-island extraction (Mean: 3.90; Difference: 0.54). This appears to be due to the fact that subject-islands were more highly rated than RC-islands for overt P-pied-piping (Mean difference = .89, compared to an average Mean difference of 0.09 for the other 3 conditions), indicating that perhaps extraction out of a subject-island overtly may be easier than out of a RC-island. Given this is unrelated to our question, however, and the result was not significant, we leave this interesting point for potential future follow-up studies.

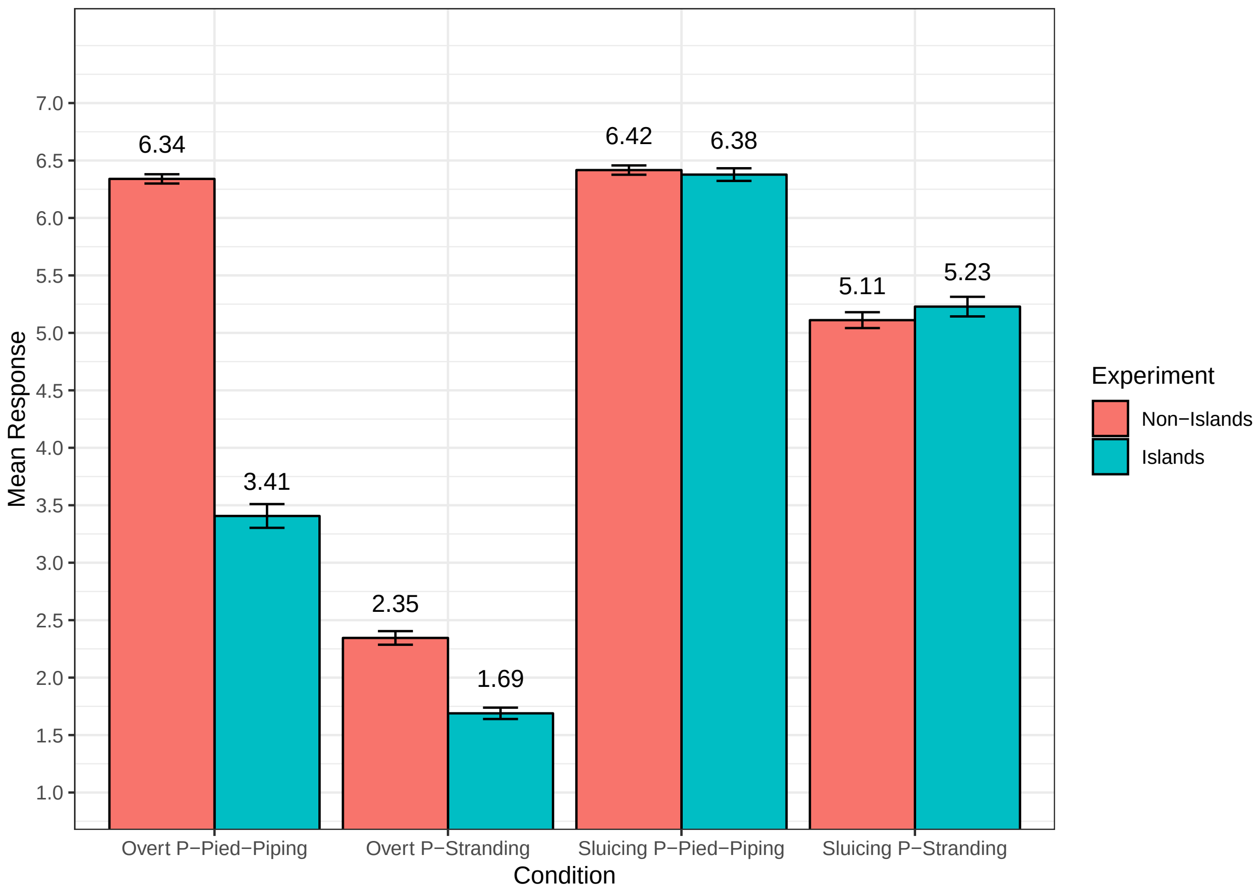
It should be noted that, as in the previous experiment, there were a number of participants (N = 12) who consistently rated P-Stranding under sluicing as less than acceptable (below 4.0), and who, once more, had the common characteristic of originating from Thessaloniki. Although again not differing in their ratings of other conditions, these participants rated P-Stranding under sluicing at 3.55 on average (Mean: 3.55; SEM: .20), as opposed to participants from other areas who rated it at 5.3 on average (Mean: 5.26; SEM: .09). Given how unequal the two sample sizes were between people originating from Thessaloniki and the rest of Greece, it would not have been statistically accurate to calculate the potential main effect of region of origin on answers. Numerically, however, the two groups appear substantially different with respect to this condition.

In order to gain more insight into the effects of islandhood on acceptability ratings both overtly and under sluicing, a comparative analysis was also run contrasting the overt and sluicing conditions from this experiment, i.e. in the context of islands, to the case-matching conditions of the previous experiment, i.e. in the context of no islands. The items used in the two experiments were as close as possible in terms of word choice, with the fillers being identical. There was also a (partial) overlap in subjects as the same pool of participants was used for both experiments. In order to run the analysis, a combined dataframe was created using all the data from Experiment 2 and all the case-matching data from Experiment 1, assigning 'Experiment' (1 vs. 2) as a between-subjects variable. Barplots showing the Mean response with SEM error bars per condition for each experiment can be seen in Figure 2.5.

An LME model was thus fitted to the combined datasets, with raw responses as the dependent variable; Sluicing (±) and P-stranding (±) as within-subjects fixed effects factors; Experiment (1 vs. 2) as a between-subjects fixed effects factor; and Subjects and Items as random effects factors, with random slopes and intercepts assumed for each. The model converged with results showing significant main effects of all fixed effects factors, each further modulated by significant two-way and three-way interactions. A significant main effect of Experiment was found ($t > 8.05; p < 8.9e - 16$), with Experiment 1 having overall higher ratings compared to Experiment 2 (Exp. 1 Mean: 5.06; SEM: 0.04; Exp. 2 Mean: 4.18; SEM: 0.06; Mean difference: 0.87). A significant main effect of Sluicing was also found ($t > 17.82; p < 1e - 16$), with sluicing conditions being on average more highly rated than their overt counterparts (Overt Mean: 3.60; SEM: 0.050; Sluicing Mean: 5.78; SEM: 0.03; Mean Difference: 2.18). The final main effect found was for P-Stranding ($t > 17.83; p < 1e - 16$), with P-pied-piping conditions being rated overall more highly than their P-stranding counterparts across both experiments (P-Pied-Piping Mean: 5.76; SEM: 0.04; P-Stranding Mean: 3.62; SEM: 0.05; Mean Difference: 2.14). All of these effects were modulated by significant two-way interactions and, more importantly, a significant three-way interaction. Specifically, a significant two-way interaction was found between P-Stranding and Sluicing ($t > 6.96; p < 3.43e - 12$), with the difference between P-pied-piping and P-Stranding conditions being significant both overtly (Mean difference: $3.05; p < .0001$) and under sluicing (Mean difference: $1.24; p < .0001$), however the size of this difference was significantly greater overtly than under sluicing across both experi-

Figure 2.5: Experiments 1 and 2: Mean Acceptability Ratings Barplot

Mean response per condition with SEM error bars. Only case-matching conditions shown from Experiment 1.



ments (Mean size difference: $1.81; p < .0001$). Furthermore, when collapsing across both experiments, the difference between overt and sluicing conditions was significant both with P-pied-piping (Mean difference: $3.08; p < .0001$) and P-stranding (Mean difference: 1.27). The model also showed a significant two-way interaction between Experiment and Sluicing ($t > 9.16; p < 1e-16$), with the difference between overt and sluicing conditions being significant both for Experiment 1 (Mean difference: $1.43; p < .0001$) and Experiment 2 (Mean difference: $3.24; p < .0001$), however the size of this difference was significantly larger for the results of Experiment 2 (Mean size difference: $1.82; p < .0001$) thanks to the amelioration of islands. Furthermore, the difference between Experiments 1 and 2 was significant only for overt conditions (Mean difference: $1.79; p < .0001$), and not sluicing conditions (Mean difference: $0.03; p > .815, n.s.$). Together with the previous findings, this lends support to the idea that, as expected, extraction out of an island overtly is significantly worse than extraction out of a non-island-containing structure; however under sluicing there is no difference between the two, with islands being ameliorated to the same degree of acceptability as structures with no island. A significant two-way interaction was also found between P-Stranding and Experiment, with the difference between P-pied-piping and P-stranding conditions - when Sluicing is not taken into account - being significant in both Experiment 1 (Mean difference: $2.65; p < .0001$) and Experiment 2 (Mean difference: $1.42; p < .0001$), but with the size of this difference being significantly larger for Experiment 1 than 2 (Mean size difference: $1.23; p < .0001$). Furthermore, although the difference between Experiments 1 and 2 was significant for P-pied-piping conditions (Mean difference: $1.49; p < .0001$) there was almost no difference between the two for P-stranding conditions (Mean difference: $0.25; p > .142, n.s.$). Finally, and most importantly, there was a significant three-way interaction found between P-Stranding, Sluicing and Experiment ($t > 6.05; p < 1.42e-09$). Ignoring the significant interactions already found and analysed above within each experiment, let us focus on the differences found when comparing the two experiments. Post-hoc comparisons taking each level of one of the variables separately from the other and checking how the the remaining two variables interacted helped clarify what the three-way interaction was due to. Specifically, when focusing on the two levels of P-pied-piping separately from each other, for P-pied-piping conditions, although the difference in acceptability judgements between the two experiments was significant for overt conditions (Mean difference: $2.94; p < .0001$), this was not true for sluicing conditions (Mean difference: $0.04; p > .995, n.s.$), with the factor of Experiment hence significantly affecting the size of this difference (Mean size difference: $2.9; p < .0001$). For P-stranding conditions, again although the difference between the two experiments was significant overtly (Mean difference: $0.64; p < .001$), it was not significant under sluicing (Mean difference: $0.09; p > .998, n.s.$), with the factor of Experiment thus significantly affecting the size of this difference (Mean size difference: $2.00; p < .0001$). Conversely, when focusing only on P-pied-piping conditions, on the one hand, although the difference between overt continuations and sluicing was not significant in Experiment 1 (Mean difference: $0.07; p > .999, n.s.$), it was significant in Experiment 2 (Mean difference: $2.97; p < .0001$), with the factor of Experiment thus having a significant effect on the size of this difference (Mean size difference: $2.89; p < .0001$). For P-stranding conditions, on the other hand, the difference between overt and sluicing condition was significant for both Experiment 1 (Mean difference: $2.78; p < .0001$) and Experiment 2 (Mean difference: $3.52; p < .0001$), with the size of this difference being significantly larger for Experiment 2 (Mean size difference: $0.74; p < .001$). Focusing now only on overt conditions, the difference in acceptability between P-pied-piping and P-stranding conditions was significant both in Experiment 1 (Mean difference: $4.01; p < .0001$) and Experiment 2 (Mean difference: $1.70; p < .0001$), however the size of this difference was significantly larger in Experiment 1 than 2 (Mean size difference: $2.31; p < .0001$). For sluicing conditions, on the other hand, although the difference between P-pied-piping and P-stranding was significant both in Experiment 1 (Mean difference: $1.30; p < .0001$) and Experiment 2 (Mean difference: $1.15; p < .0001$), the size of this difference did not vary significantly from one experiment to the other (Mean size difference: $0.15; p > .167, n.s.$).

Together all these results contribute to the conclusion that a) extracting out of an island overtly, on the one hand, is significantly less acceptable than extracting out of a non-island-containing source; b) extracting out of an island under sluicing, on the other hand, is as acceptable as extracting out of a non-island-containing source; c) both (a) and (b) are true for both P-pied-piping and P-stranding; d) regardless of whether we are extracting out of an island or not, P-stranding becomes significantly more acceptable under sluicing than it is overtly; e) P-stranding out of an island overtly appears to have a compounding effect compared to P-pied-piping out of an island and P-stranding out of a non-island-containing structure; and finally f) regardless of whether we are extracting out of an island or not and whether we are looking at sluicing or non-sluicing conditions, P-pied-piping is significantly more acceptable than P-stranding.

Following the same method as previously, the conditional R^2 of the comparative model was calculated at 0.85, meaning that the model as is (with random effects factors) accounted for approximately 85% of the variability in the combined data. A power analysis (following Westfall et al. (2014)) estimated the study power at 1, meaning the same result would be obtained 100% of the time (i.e. a very robust result).

2.2.5 Discussion

In summary, from the results of Exp. 2, as well as their comparison with those of Exp. 1, we can see that overtly extracting a PP out of an island, whether of the RC or subject kind, is unacceptable compared to P-pied-piping with no island. This result was expected based on any sluicing account. Furthermore, even though in both experiments P-stranding was much less acceptable overtly than P-pied-piping, the combination of island extraction with P-stranding in Exp. 2 makes overt P-stranding ratings significantly worse, indicating that the two overt violations likely have an additive effect on acceptability. Although this result, in particular, may be in line with a processing approach advocating for additive processing costs affecting acceptability ratings in both overt and sluicing constructions, it is not in line with any of the purely theoretical approaches to sluicing we have examined thus far, as all of these approaches predict that regardless of the number of violations, all unacceptable structures should receive uniform floor ratings.

As also expected by all accounts, island extraction was ameliorated under sluicing, with P-pied-piping out of an island being significantly better under sluicing than overtly. Unexpectedly for a PF-Deletion Hypothesis, however, as in Exp. 1, the same appears to be true for P-stranding under sluicing. Furthermore, even though P-stranding was significantly better under sluicing than overtly, it was still not as acceptable as P-pied-piping under sluicing, with a significant difference between the two, as in Exp. 1. Moreover, as shown by a significant three-way interaction of the factors of Sluicing, P-stranding and Experiment, even though there was a large difference in both P-pied-piping and P-stranding conditions overtly between the two experiments, there was no significant difference - no difference whatsoever, in fact - between the two experiments when it came to sluicing conditions. This last point, in particular, is important for several reasons.

Firstly, this is the first experiment of its kind to finally provide evidence that the phenomenon of island amelioration under sluicing does indeed result in ratings absolutely equal to a) sluicing involving no islands in the antecedent; b) overt structure with no islands or other violations. Although this has been one of the most (in)famous purported characteristics of sluicing ever since its genesis, no experimental evidence has since confirmed it practically and with decimal precision accuracy. Secondly, having this acceptability profile for island amelioration under sluicing in Greek as a baseline, we can now compare the results of P-stranding amelioration to this baseline in order to ascertain whether they are due to PPs simply being a form of PF-island in Greek. Specifically, as predicted in section 2.2.1, if a) PPs are simply some form of PF-island to be ameliorated under sluicing similarly to other islands, with amelioration under sluicing being due to ungrammatical features not being pronounced; and b) the deterioration in acceptability between P-stranding and P-pied-piping found under sluicing in Exp. 1 was due to the cost of this amelioration process, then we would expect a specific set of results. Firstly, P-stranding and P-pied-piping under sluicing in this experiment should have been equal to each other, both involving the amelioration of island(s). Secondly, both these conditions should be equal to P-stranding under sluicing in Exp. 1, given all three involve island amelioration. Finally, all three of these conditions should be significantly less acceptable than P-pied-piping under sluicing in Exp. 1, given the latter involves no island amelioration. However, as seen above, we found no evidence for any of these predictions.

The results of both these experiments, in isolation and in combination, are not explained by any of the theoretical syntax approaches we have thus far seen. With respect to a pointer account, as also mentioned in the discussion section of Exp. 1, its major faltering point is in explaining the significant difference between P-stranding and P-pied-piping under sluicing if there is no structure assumed within the *e*-site. With respect to Nykiel (2013)'s pointer account, more specifically, with its concept of remnant and correlate complexity as the driving force behind sluice ratings, these gradient results also remain uncaptured; this is particularly true when we take into consideration that in Exp. 1 there was a significant main effect of remnant complexity, but in the opposite direction of that predicted by a pointer account (i.e. with *wh*-remnants being slightly more acceptable than *which*-NP remnants), however, more importantly, there was no interaction of this factor with either P-stranding or Sluicing. This factor also did not present with a main effect nor significant two- or three-way interactions in Exp. 2.

In conclusion, although various aspects of both experiments are as expected by all accounts, such as P-stranding overtly with no island being unacceptable, and overt extraction out of an island also being unacceptable, no theory we have encountered thus far can actually account for the full set of data from these experiments, from the purely theoretical PF-Deletion or LF-Copying to the more processing-based pointer accounts. To these two apparently perplexing sets of data, we will now add another which further investigates previously proposed ideas on P-stranding allowance under sluicing and underlines their explanatory inadequacy with respect to this full set of results.

2.3 Experiment 3: P-Stranding under Contrast Sluicing

2.3.1 Experimental Outline, Aim and Predictions

Having seen that island amelioration does not have any effect on acceptability ratings under sluicing, but also that it seems to behave differently to regular P-stranding 'amelioration' under sluicing as witnessed in Experiment 1, we come to Experiment 3 which is designed along the same lines as the previous two, though on a smaller scale, this time examining a special case of sluicing purported to have a unique relationship with island amelioration. Its focus, specifically, is on P-Stranding under contrastive sluicing. Contrastive focus in particular appears to have a much more restrictive effect on sluicing, simultaneously disallowing both island amelioration, but also a cleft or copular source. The motivation for using this type of sluicing is hence two-fold.

On the one hand, according to Griffiths and Lipták (2014) (see also M. Barros et al.

(2014); Merchant (2008), though cf. Culicover and Jackendoff (2005); Jacobson (2016); Weir (2014)), repair of otherwise ungrammatical structures under IP-ellipsis is related to the presence or not of contrastive focus, as in for instance (62) vs. (63), with contrastive focus represented by capitalisation [15]. Based on such comparisons, they present an updated version of island repair under ellipsis, shown in (64).

(62) Hannah wanted to find someone who can make a special sort of pie, but I don't remember what sort (of pie) [~~she wanted to find someone who can make *t*~~].

(63) * Hannah wanted to find someone who can make PUMPKIN PIE, but I don't remember what OTHER PIE [~~she wanted to find someone who can make *t*~~].

(64) **Generalisation on island repair**
'Contrastive fragments cannot repair islands. Non-contrastive fragments can potentially repair islands.'

(Griffiths & Lipták, 2014, p. 32, (110))

If one were, thus, to equate P-stranding amelioration under sluicing to a form of island repair, as we have seen being argued, one would expect to also find no P-stranding allowance under constrast sluicing, even if this is supposedly possible under regular sluicing[16]. Hence, if this is the true reason for the effects we found in the previous two experiments, we would expect, in contrast, to find P-stranding disallowed both overtly and under sluicing here. With respect to the predictions of the PF-Deletion hypothesis more generally, i.e. without PPs behaving as PF-islands, these predictions do not change for contrast or 'else' modification sluices (Merchant, 2001), i.e. the PSG is once again expected to hold in this case of sluicing. The predictions of the pointer accounts also do not change compared to the previous experiments. To reiterate, a general pointer account would predict P-stranding to be acceptable under sluicing to the same degree that P-pied-piping is; and a pointer account hinging on remnant/correlate complexity would predict such acceptability under sluicing to be correlated with this degree of complexity, with more complexity of the correlate affording higher acceptability of P-stranding under sluicing and less complexity lower acceptability.

On the other hand, a second, very important motivation for examining P-stranding under contrast sluicing is in order to gain more confidence in ruling out the possibility of an alternative, copular source underlying apparent P-stranding under sluicing in Greek. The reasoning in this case is that contrast sluices or 'else' modifications are technically disallowed

[15] Judgements based on Merchant (2008).
[16] NB: I do not present judgements at this time on regular island repair under contrast sluicing in Greek, as I have not yet run properly controlled acceptability rating studies and do not wish to draw conclusions from the few judgements of myself and friends/family, although anecdotally I can say that regular islands were indeed considerably less acceptable than P-stranding. Due to temporal constraints, these studies will be conducted in due course.

with copulas cross-linguistically, and I would argue this to also be true in Greek (65). That is to say, even if we were to ignore the fact that a cleft source remnant should appear in Nominative, and not Accusative as the remnants in this and the previous experiment do, if a cleft or copular structure is indeed the true source of Greek P-stranding in the previous two studies, then we should find P-stranding to be unacceptable in this study.

(65) a. *I Maria chorepse me to GIANNI, alla de
Det.F.NOM *Maria*.F.NOM *danced*.3SG *with* *Det*.M.ACC *Giannis*.M.ACC *but* NEG
thimame pjos ALLOS itan.
remember.1SG *who*.M.NOM *else*.M.NOM *was*.3SG
'*Maria danced with GIANNIS, but I do not remember who ELSE it was.'

b. *I Maria chorepse me pende NEAROUS, alla de
Det.F.NOM *Maria*.F.NOM *danced*.3SG *with* *five* *young.men*.M.ACC *but* NEG
thimame poses KOPELES itan.
remember.1SG *how.many*.F.ACC.PL. *young.ladies*.F.ACC.PL. *was*.3SG
'*Maria danced with five GUYS, but I do not remember how many GIRLS it was.'

2.3.2 Methods

A web-based acceptability judgement task was performed as in Experiments 1 and 2 using the online platform of Ibex Farm®, with participants recruited via Facebook®.

The factors of Sluicing (±) and P-Stranding (±) were once more crossed, this time in the environment of contrast sluicing. This resulted in the 4 conditions presented in Table 2.17.

In order to attempt to convey contrastive focus to the same degree in all items and conditions, the stimuli in this study were much more tightly controlled in their form than those in the previous two experiments, absolutely following the design presented in the next section, with no embellishments. A total of 40 experimental stimuli were presented to each subject, 10 of each condition. Out of the 40 native speakers recruited, for the same reasons as in the previous Experiments, the data of only 32 were used, giving 320 data points per condition.

2.3.3 Materials

The experimental stimuli were based on those of the two previous experiments, using the same lexical items where possible and adhering to the same criteria as in those, again where possible (see below), in order to generate new stimuli where necessary. The total number of experimental stimuli was 40, complemented by 80 fillers (with the same distribution as

Table 2.17: Experiment 3 Conditions and Examples

(66) I fititria singatiki apo perisi me ton Gerasimo, alla de thimame . . .
*Det.*F.NOM *student.*F.NOM *cohabits.*3SG *from last.year with Det.*M.ACC *Gerasimos.*M.ACC *but* NEG *remember.*1SG
'The student has been living with Gerasimos since last year, but I don't remember . . . '

	Condition	Example
1	Non-Sluicing, P-Pied-Piping	. . . me pjon allon singatiki. *with who.*M.ACC *other.*M.ACC *cohabits.*3SG '. . . with who else she has been living.'
2	Non-Sluicing, P-Stranding	. . . pjon allon singatiki me. *who.*M.ACC *other.*M.ACC *cohabits.*3SG *with* '. . . who else she has been living with.'
3	Sluicing, P-Pied-Piping	. . . me pjon allon. *with who.*M.ACC *other.*M.ACC '. . . with who else.'
4	Sluicing, P-Stranding	. . . pjon allon. *who.*M.ACC *other.*M.ACC '. . . who else.'

in the previous experiments). As mentioned above, to control for the way contrastive focus was conveyed, stimulus design was based exclusively on the pattern presented in (67), e.g. (68).

(67) Main Clause Subject (Feminine descriptive NP) + Main Verb + Adverb + Preposition + Internal Argument (Masculine proper name), 'but' (+ Second Clause Subject) + NEGATION + Embedding Verb + Interrogative Pronoun (Accusative) + ELSE.

(68) I epimelitria paraponiotan sinechos jia to
Det.F.NOM *class.prefect*.F.NOM *was.complaining*.3SG *constantly for Det*.M.ACC
Mano, alla den akousa (jia) pjon allon.
Manos.M.ACC *but* NEG *heard*.1SG *(for) who*.M.ACC *else*.M.ACC
'The class prefect was constantly complaining about Manos, but I didn't hear (about) who else.'

Furthermore, given the nature of the study did not allow for contrastive focus to be phonologically indicated, and presenting one word in all caps or bold font was considered too visually distracting, the way that contrastive focus was conveyed was by establishing the intended focused correlate as the *only* discourse focused element. To achieve maximum discourse focus on this particular element, items were designed with the correlate being the only proper name in the main clause (e.g. *Manos*), with a gender-unambiguous feminine descriptive NP used as the subject of the main clause (for an overview of the relative effectiveness of proper vs. descriptive names in establishing discourse focus, see Sanford, Moar, and Garrod (1988)). An AdvP was added after the main verb to increase the naturalness of the stimuli.

With respect to the various other factors which had been controlled for in the previous two experiments, given all stimuli adhered strictly to the design in (67), with a singular proper masculine name as P argument, it was not possible to check for differences between singular and plural, nor simplex and complex remnants. Half the embedded clauses, however, were introduced by a verb which patterned with Accusative and the other half not, in order to check for grammatical illusion effects.

With respect to fillers, the shortest fillers of the previous two studies were used (Mean: 17.67; SD: 1.77) and were made up of simple active, passive, DP-ellipsis and VP-ellipsis structures. There were 80 in total, split between 40 grammatical and 40 ungrammatical ones. As in the previous two studies, comprehension questions were also included for 2/3 of non-sluicing and sluicing P-pied-piping stimuli, and half of the grammatical fillers.

2.3.4 Results

As in the previous two experiments, data was cleaned following the same exclusion criteria, leaving the data of 32 participants. Furthermore, outliers calculated at 2.5*SDs from the Mean per condition were Windsorized. There were no responses measured below $100ms$, leaving us with 1,200 data points in total, or 320 per condition.

A maximal LME model was successfully fitted to the data using R's lme4 package, with raw responses as the dependent variable, the two fixed effects factors of P-stranding ($\pm$) and Sluicing ($\pm$) and random effects factors of subjects and items included (random slopes and intercepts assumed for each), using the standard optimizer. As with previous experiments, although the data were skewed and not normally distributed, this was not taken as a counter-indication to the accuracy of the LME model. Once again, however, in order to have confidence in the model's results, different LME models were also fitted with $Log10$ transformations, z-scores and $Log10$ transformed $z-$scores of raw responses as the dependent variable, but also using different optimizers (*optimx*), each of which could have given a better-fitting model, had non-normality been a significant issue. As the fit of these additional models was not significantly different to that of the original model using the dependent variable of raw answers, with only small differences in effect sizes, only the latter is presented in-text here, with the rest presented in Table 2.19.

As in the previous two experiments, results showed significant main effects for both the fixed factors of Sluicing and P-Stranding; with respect to the factor of Sluicing ($t > 4.625; p < 6.661e-16$), sluicing conditions were rated as overall more acceptable than overt continuation conditions (Mean difference: .86); and with respect to the factor of P-Stranding ($t > 12.24; p < 1e-16$), this was due to P-pied-piped conditions being rated on average as significantly more acceptable than their P-stranded counterparts (Mean difference: 2.85). Importantly, there was also a significant two-way interaction between P-Stranding and Sluicing ($t > 7.95; p < 4.218e-15$), with post-hoc Tukey-adjusted comparisons showing that this was due to the difference between P-stranding and P-pied-piping conditions being significant both overtly (Mean difference: $4.23; p < .0001$) and under sluicing (Mean difference: $1.50; p < .0001$), however the size of this difference was significantly greater overtly than under sluicing (Mean size difference: $2.73; p < .0001$). Furthermore, the difference between overt and sluicing conditions was significant when paired with P-stranding (Mean difference: $2.26; p < .0001$), but not when paired with P-pied-piping (Mean difference: $0.47; p > .229, n.s.$), with the factor of P-Stranding thus having a significant effect on the size of this difference (Mean size difference: $1.79; p < .0001$). These results show that, as expected from the results of the previous two studies, P-stranding in the context of contrastive focus becomes significantly better under sluicing than it is overtly, with no such difference in acceptability ratings for P-pied-piping conditions; furthermore, despite this amelioration under sluicing, P-stranding is still significantly less acceptable than P-pied-piping both overtly and under sluicing.

As with the previous experiments, the results are better translated into Mean barplots with SEM error bars, as in Figure 2.6.

Figure 2.6: Experiment 3: Mean Acceptability Ratings Barplots

Explanatory Notes: Mean response per condition with SEM error bars

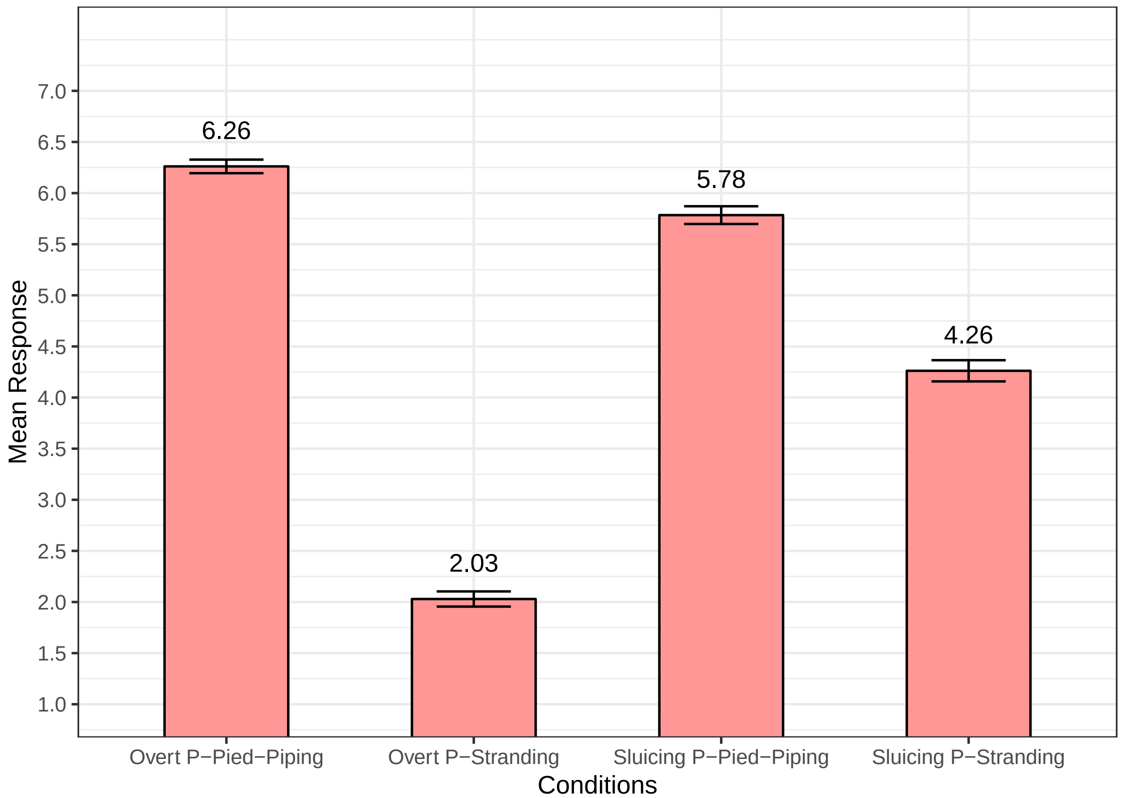


Table 2.18: Experiment 3: Acceptability Ratings Numerical Summary

	Condition	Response	SD	SEM	95% CI
1	Non-Sluicing P-Pied-Piping	6.261	1.185	0.066	0.131
2	Non-Sluicing P-Stranding	2.029	1.327	0.074	0.146
3	Sluicing P-Pied-Piping	5.784	1.547	0.087	0.171
4	Sluicing P-Stranding	4.262	1.857	0.104	0.204

With respect to the additional factor of Embedding Cerb (±Acc.), its inclusion in the model provided no significant main effect nor interactions with any other fixed effects factor

Table 2.19: Experiment 3: Main Effects & Interactions

		Raw Scores		z-Scores		$Log10$-Scores	
	Factor	t-value	p-value	t-value	p-value	t-value	p-value
1	P-Stranding	12.2356	< 0.0001	17.144	< 0.0001	10.973	< 0.0001
2	Sluicing	4.6254	< 0.001	4.707	< 0.001	5.834	< 0.001
3	P-Stranding*Sluicing	7.9523	< 0.0001	8.518	< 0.0001	8.532	< 0.0001

(all $t's < 0.6; p's > .55, n.s.$), indicating that overall there were no grammatical illusion effects due to the c-selection properties of the embedding verb. With respect to the factor of PP Type, its inclusion in the model produced a marginally significant three-way interaction between P-Stranding, Sluicing and PP Type ($t > 2.03; p < .042$). This was due to the two Ps 'me' (*with*) and 'se' (*to*) showing significantly more acceptable P-pied-piping under Sluicing (Mean for 'me': $6.25; SD: 1.18$; Mean for 'se': $6.27; SD: 1.00$) compared to the two Ps 'jia' (*for*) and 'apo' (*from*) (Mean for 'jia':$5.56; SD: 1.68$; Mean for 'apo': $5.38; SD: 1.76$), with no such difference for the other conditions (all other differences ranged between $.06 - .35$). Given this effect is marginal, it is unclear without further replication attempts whether it is reliable or not. If it is, then it is also unclear what the cause of this effect is, i.e. whether it is due to 'me' and 'se' being slightly shorter in length than 'apo' and 'jia' or some other, as yet unknown, semantic factor related to the way contrastive focus is achieved (since this effect was not present in either of the previous studies). As this concerned only P-pied-piping under sluicing and not P-stranding, I will simply observe this interesting finding and leave it for further confirmatory testing in the future without considering it strictly relevant to the story at hand.

As an indicator of the overall model fit, as in the previous studies we can calculate the conditional R^2 for the GLME model based on Nakagawa and Schielzeth (2013). This gives us a conditional R^2 of .82, meaning that this model, as presented, i.e. including both fixed and random effects factors, can account for approximately 82% of the variation in the data (a large effect size, following Sullivan and Feinn (2012) and adapted from Ferguson (2009)). A power analysis (following Westfall et al. (2014)) estimated the study power at 1, meaning the same result would be obtained 100% of the time (i.e. a very robust result).

It should be noted, once more, that the factor of region of origin did appear to play a role in the results of this experiment, though not one possible to calculate with statistical accuracy. All but 2 of the recorded Northerners from the region of Thessaloniki (Total N = 8) consistently marked P-Stranding under sluicing as less acceptable than did the other participants (North: Mean: $3.43 SEM: .13$, South: $4.23, SEM: .11$). These results follow the same trend observed in the previous two experiments. This shall be examined in more detail in section 2.4.

2.3.5 Discussion

In summary, the results of this study show that even in the context of contrast sluicing P-stranding is significantly more acceptable under sluicing than it is overtly. This is not something predicted by a PF-Deletion account, nor should it be possible if PPs are simply another island which is generally ameliorated under sluicing for Greek, given islands are generally not ameliorated under contrast sluicing. As with the results of the previous studies, despite this amelioration of P-stranding under sluicing, the acceptability difference between P-pied-piping and P-stranding under sluicing is also not successfully captured by either an LF-Copying account, nor a pointer account.

Interestingly, when comparing these results to those of the previous two experiments, the results shown here for contrast sluices are overall worse than either their regular sluice or island-extracted counterparts from Experiments 1 and 2 respectively, both of which had shown similar acceptability ratings, with P-pied-piping sluicing having a Mean of 6.3 and P-stranding sluicing a Mean of 5.15 on average across both studies. Here, the acceptability scores are much lower for both types of sluicing, with P-pied-piping sluices scoring a Mean of 5.74 and P-stranding sluices a Mean of 4.23. Given contrast sluicing seems to have thus affected both levels of P-stranding approximately equally (around 0.8 points difference between the Means of Experiments 1 and 2 on average, on the one hand, and 3, on the other), I am inclined to believe that this is not due to some inherent effect that contrastive focus appears to have on P-stranding amelioration patterns per se (e.g. following Griffiths and Lipták (2014)), but rather that this difference is more likely to be a result of the study modality and design in Experiment 3 simply not conveying the necessary intended contrast strongly enough, despite our best efforts. It would thus be interesting to see if a more strongly expressed contrast via the more common phonological means or through orthographic capitalisation and/or use of different font would reduce this intra-experimental variation in sluicing results.

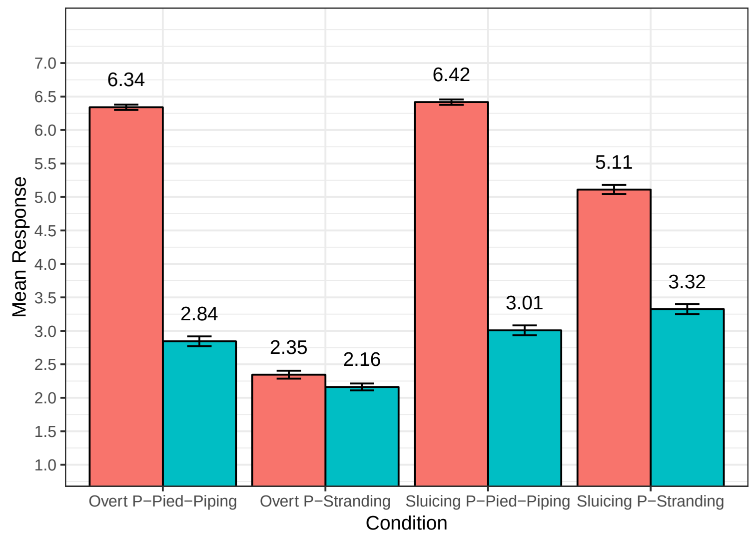
2.4 General Discussion for all Greek Acceptability Judgement Studies

The results of all three of these experiments (repeated altogether below) show that, without a doubt, Greek has the makings of another PSG-defying language, at least for the majority of its dialects, with potentially the exception of one.

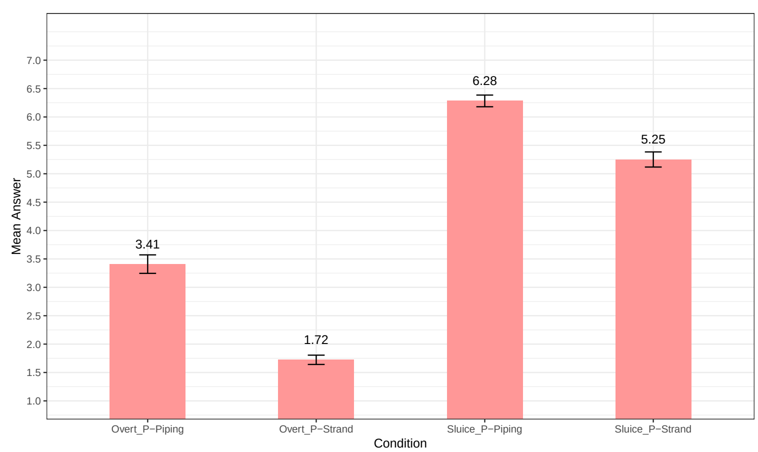
Specifically, Greek appears to allow P-stranding under regular sluicing (Experiment 1), sluicing out of supposed islands (Experiment 2) and contrast sluicing (Experiment 3). Across all three experiments, P-stranding under sluicing appeared significantly more acceptable than overtly, but also slightly less acceptable than P-pied-piping. For regular sluicing and

Figure 2.7: All Greek Acceptability Results

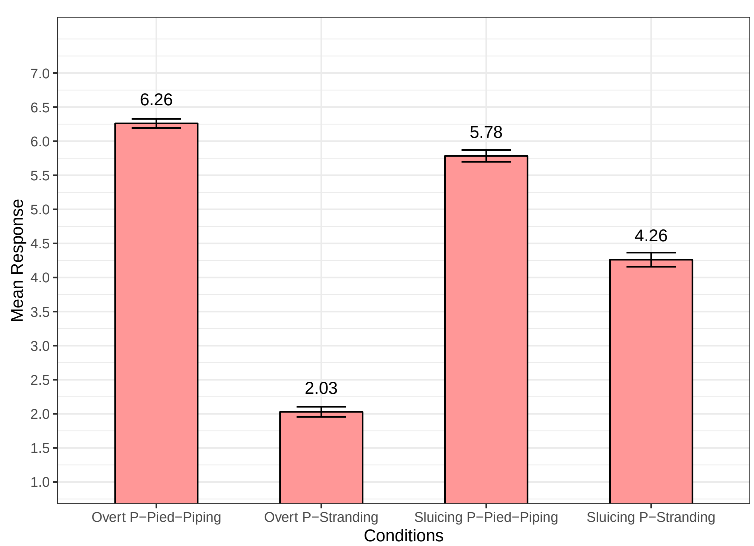
(a) Experiment 1 (Case-Matching*P-Stranding*Sluicing)



(b) Experiment 2 (P-Stranding*Sluicing in context of Islands)



(c) Experiment 3 (P-Stranding*Sluicing in Contrast Sluicing)



sluicing out of supposed islands, both these conditions under sluicing were almost identical between the two studies, with contrast sluicing results being generally slightly lower. This results pattern may be due to the study modality of Experiment 3 not being particularly well-suited to contrastive focus or it could simply be circumstantial. For Experiments 1 and 2, a between-subjects analysis confirmed that there was no significant difference between the sluicing results of the two studies. It is interesting to note that the 'presence' or not of an island within the *e*-site did not have any effect on sluicing results for P-pied-piping, with the results of this condition being equal across both studies and significantly different from those of P-stranding. This would perhaps indicate that there is something different occurring with P-stranding in comparison to general island amelioration, at least insofar as RC- and subject-islands are concerned. The results of the third experiment, in particular, further add to this idea that P-stranding under sluicing differs from island amelioration, thanks to P-stranding also being rated as acceptable under contrast sluicing in this context, as opposed to other PF-islands Griffiths and Lipták (2014). That having been said, as also mentioned previously, Greek island behaviour does still remain to be properly confirmed in the context of contrast sluicing in follow-up studies.

Furthermore, with respect to the predictions as laid out for each experiment, we were able to see that a PF-Deletion account as it stands, with or without the Fit Condition amendment, is not a good fit for the data, nor is one along the lines of an LF-Copying account. Even a pointer account, with or without emphasis on remnant complexity, is unable to capture the results of these three experiments. This leaves us with the question of how exactly to explain these data in full, but also of where our Greek results stand with respect to those of other languages. Specifically, the lack of controlled large-scale experiments in other languages, apparently PSG-defying or not, means that we do not know if Greek is somehow special, allowing P-stranding under sluicing despite no clear grammatical source being available; or whether this pattern we have observed is actually typical of all overtly non-P-stranding languages, with it simply being the case that no one has adequately checked before. This lack of prior experiments also means that we do not know exactly how Greek results compare to those of PSG-defying languages which have alternative grammatical sources available to them for apparent P-stranding under sluicing. The one larger dataset we do have is from Saudi Arabic, thanks to the work of AlShaalan and Abels (2019). The results from that study indicate that when there is an alternative source available, there is no acceptability difference between P-stranding and P-pied-piping under sluicing. This is taken as indication for the alternative source being chosen. However, when this alternative is blocked, P-stranding is still - perplexingly - significantly more acceptable under sluicing, albeit less so than P-pied-piping, a situation reminiscent of the pattern also found here. This leaves us with an even more burning question of whether there is an elegant way to explain all these data together. Before attempting to answer that question, further data must first be gathered in order to ascertain the generalisability or not of these Greek findings to other non-P-stranding languages. This shall be the focus of the next Experiment.

Before moving on to the next section, however, there are a couple of points concerning the Greek dataset which must be addressed. Firstly, in the above results we saw that Greek allows P-stranding acceptably under sluicing for PP-arguments, however we did not investigate whether this was also true for PP-adjuncts, leaving this for future investigation. Anecdotally, I can attest to my own judgements and those of at least 10 other native Greek speakers (one linguist), that P-stranding does appear to also be acceptable for PP-adjuncts, however further testing is, of course, required.

Another point which deserves to be briefly expanded upon is the outlying data of Northern Greek speakers with respect to P-stranding. In particular, on average across all three studies, 9 out of 10 of the subjects that did not appear to accept P-stranding under sluicing originated from the North of Greece and more specifically the area in and around Thessaloniki[17]. Although time constraints have kept me from extensively testing out the below theory experimentally, it may be possible to explain this variability in a way that allows all Greek data to be unified. This explanation is based on two points; the first is the widely-attested Greek dialectal variation with respect to the expression of an indirect object with either a full PP or a bare argument; and the second is how Greek sluicing appears (at least anecdotally) to allow for an alternation in these two expressions between remnant and correlate.

With respect to the first point, although much variability outside phonology has unfortunately been lost between Greek dialects due to government-mandated Greek lessons in school, there is one notable syntactic difference between Northerners and the rest of Greece in how the cases of Genitive and Accusative can be used in the expression of indirect arguments. Modern Greek has 4 cases (Nominative, Genitive, Accusative and Vocative), however Ancient Greek additionally utilised Dative case to express indirect objects in double object frames (among other things), with this use of case being taken over by the other, modern cases over time. Importantly, this diachronic syncretism appears to have varied with dialect, with Dative being absorbed by Genitive in all other areas of Greece, and by Accusative in the Northern dialect (see also Joseph, Philippaki-Warburton, and Philippaki-Warburton (1987)). Indeed, this is one of the main linguistic features which can reliably identify native speakers from Thessaloniki. In addition to these two cases, it is possible to express an indirect object as a full PP (P + NP.ACC) cross-dialectally. As a result, if one were to create a Dative alternation pattern for Greek, it would look something like the one in Table 2.20.

Moving on to the second point of our explanation, we come to what I will term 'Dative al-

[17]It is also worthwhile noting here that all judgements recorded in Merchant's 2001 work, and subsequently, with respect to sluicing also came from Northern speakers (Merchant, 2016, p.c.; Giannakidou, 2017, p.c.).

Table 2.20: Greek Dative Alternation pattern by dialect

	Dialect	Representation of Indirect Argument	Example
1	South	P + Acc./ bare Gen.	I.give to you.ACC the book.ACC. / You.GEN I.give the book.ACC.
2	North	P + Acc./ bare Acc.	I.give to you.ACC. the book.ACC. / You.ACC. I.give the book.ACC.

ternation under sluicing' for Greek[18]. As just mentioned, it is possible to express an indirect argument in Greek either within a double object (DO) frame or an NP + PP frame. These two frames alternate freely in certain environments, although not all (see Appendix, section B.4, for further details on this alternation in Greek), as is also the case in English, with considerable literature existing - at least insofar as English is concerned - on the identity or not of the two structures (Beck and Johnson (2004); Green (1974); Kayne (1984), a.m.o.). Intriguingly, and subject to further testing, it would appear as though Greek sluicing also allows this alternation between correlate and remnant, i.e. it is acceptable to have one form appear in the correlate with the other form appearing in the remnant, with either order (PP correlate - DO remnant; DO correlate - PP remnant) being acceptable. This effectively amounts to case-mismatching at least as far as the original case-matching generalisation is concerned, although not with the Fit Condition amendment (Abels, 2016). The one caveat I have observed to this alternation is that its acceptability appears to vary depending on the amount of material present between correlate and remnant, with more material lending greater acceptability, leading to the pattern in (69).

(69) a. */?? O Markos edose to fakelo se
Det.M.NOM *Markos*.M.NOM *gave*.3SG *Det*.M.ACC *folder*.M.ACC *to*
kapjon, ala de thimame tinos.
someone.M.ACC *but* NEG *remember*.1SG *who*.M.GEN
'Marcus gave the folder to someone, but I do not remember who.'

b. O Markos edose to fakelo se kapjon
Det.M.NOM *Markos*.M.NOM *gave*.3SG *Det*.M.ACC *folder*.M.ACC *to someone*.M.ACC
chthes arja to vradi, ala de thimame tinos.
yesterday late Det.N.ACC *night*.N.ACC *but* NEG *remember*.1SG *who*.M.GEN
'Marcus gave the folder to someone late last night, but I do not remember who.'

It is unclear exactly why this structurally irrelevant material should make such a difference, however one possibility is that it acts as some form of parsing 'buffer' between remnant

[18] This term was chosen despite the absence of Dative per se in Modern Greek in order to be in keeping with similar phenomena in other languages, e.g. Icelandic (Wood et al., 2016).

and correlate; due to our working memory limitations, the further the remnant gets from the correlate, the less the exact form of the correlate may be primed or may matter, allowing for its structurally equivalent alternative to appear in its stead, both of which are acceptable with a full overt continuation. When this alternative is otherwise blocked, however, such as in locative structures (70), the alternation under sluicing is also disallowed, regardless of additional material (71).

(70) a. O Markos petakse to fakelo se enan potamo.
*Det.*M.NOM *Markos.*M.NOM *threw.*3SG. *Det.*M.ACC *folder.*M.ACC *to one.*M.ACC *river.*M.ACC
'Marcus threw the folder into a river'

b. # O Markos petakse enos potamou to fakelo.
*Det.*M.NOM *Markos.*M.NOM *threw.*3SG. *one.*M.GEN *river.*M.GEN *Det.*M.ACC *folder.*M.ACC.
'Marcus threw the folder into a river'

(71) a. * O Markos petakse to fakelo se kapjon potamo, ala de thimame tinos (potamou).
*Det.*M.NOM *Markos.*M.NOM *threw.*3SG *Det.*M.ACC *folder.*M.ACC *to some.*M.ACC *river.*M.ACC *but* NEG *remember.*1SG *which.*M.GEN *(river.*M.GEN*)*
'Marcus threw the folder into some river, but I do not remember which (river)'.

b. * O Markos petakse to fakelo se kapjon potamo chthes arja to vradi, ala de thimame tinos (potamou).
*Det.*M.NOM *Markos.*M.NOM *threw.*3SG *Det.*M.ACC *folder.*M.ACC *to some.*M.ACC *river.*M.ACC *yesterday late Det.*N.ACC *night.*N.ACC *but* NEG *remember.*1SG *which.*M.GEN *(river.*M.GEN*)*
'Marcus threw the folder into some river late last night, but I do not remember which (river).'

The combination of these two points may yield an explanation for the Northern P-stranding data. Specifically, given the availability of a bare Accusative-marked remnant to potentially serve as a full PP, what may be occurring when Northerners encounter a 'P-stranded' remnant is that they are interpreting it as a full PP and not a case of P-stranding, i.e. in the same way that Southerners would a bare Genitive-marked remnant. If this is the case, then two predictions can be made; firstly, adding more material between correlate and remnant may lead to this remnant being *acceptably* treated as a Dative alternative to the correlate PP; and secondly, when such a Dative alternation is a priori excluded as a possibility, such as with locative constructions, the Accusative-marked remnant should be treated instead as a P-stranded remnant and be considered more acceptable. Initial testing with a handful of speakers has shown these predictions to be accurate. That is to say, the same Northern speakers that found P-stranding unacceptable found it more acceptable when more material was added - in the words of one informant 'The more material you add, the better it sounds' - as well as when a locative structure was used as in (71)[19], regardless of the presence of additional material. Of course, all of these points are tentative and require further testing to be clarified. However, the explanation I shall provide in the next chapter for all the data presented thus far can also account for this particular dialectal behaviour.

2.5 Experiment 4: P-Stranding under Sluicing in German

2.5.1 Experimental Outline, Aim and Predictions

In the wake of the Greek acceptability studies, the question remains how and why Greek in particular appears to defy the PSG, despite no grammatical alternative being possible. As mentioned in the previous section, one major outstanding query concerns whether our results are interesting due to some, as yet undefined, specific aspect of Greek - and perhaps other similar languages - or whether they are remarkable simply because they were obtained through the first appropriately controlled and powered study of its kind. That is to say, it may be the case that every other supposedly PSG-compliant language will also show similar results when put to the test, with P-stranding scoring considerably higher under sluicing compared to overtly, but still worse overall compared to P-pied-piping. If we are to develop an appropriate theory to explain these data, we need to first identify how generalisable these results are.

To test this, I decided to run a replication study of the first Greek experiment in German, a language which - like Greek - shows strong, overt feature-marking and is therefore easy to check for case-matching. It is also similar in that it, too, has been repeatedly argued in the literature to be fully compliant with both form-identity generalisations (Merchant, 2001; Van Craenenbroeck, 2012), although no experiments have thus far been conducted to practically confirm this. As in Experiment 1, therefore, the factors of P-Stranding (±), Case-matching (±) and Sluicing (±) were crossed in German. The results of this study would clarify whether both form-identity generalisations hold in German, as has been argued, or

[19] Of course, as expected, a Genitive-marked bare remnant was completely unacceptable for them and remained so irrespective of the amount of material between itself and the correlate.

whether it too, like Greek, would show an intriguing variety in its results, particularly with respect to the acceptability of P-stranding.

The sets of predictions outlined here are the same as those made for the original Greek study. That is to say, under a PF-Deletion Hypothesis, with or without the Fit Amendment, one would predict that P-stranding under sluicing in German should be as unacceptable as it is overtly, with no interactions predicted with either the factor of Sluicing or Case-matching. Similar predictions can be made for the factor of Case-matching, with a significant main effect predicted, and no interaction with Sluicing. In other words, both P-stranding and case-mismatching should always be unacceptable and at floor level, regardless of the presence or not of a sluicing environment. We should, thus, also not expect a main effect of Sluicing. With respect to the interaction of all three factors, there is no reason to predict a compounding effect for the two grammatical violations of P-stranding and case-mismatching when present together either overtly or under sluicing, given that when we have either P-stranding or case-mismatching or both, a PF-Deletion Hypothesis would predict equally unacceptable ratings at floor level. As such, we would not anticipate a significant three-way interaction between P-Stranding, Case-matching and Sluicing in any direction.

Moving on to a slightly different version of this PF-Deletion Hypothesis, if we assume that PPs are PF-islands in German, regardless of the fact that there was no support for this theory from Experiments 2 and 3 in Greek, there would be a slight difference in predictions, particularly with respect to the factor of P-Stranding. Starting first with the factor of Case-matching, all predictions involving this factor should remain the same as for a regular PF-Deletion Hypothesis, i.e. regardless of the presence or not of Sluicing or P-Stranding, case-mismatching should always be rated as completely unacceptable and at floor level. As such, one would predict a main effect of Case-matching and no interaction between Case-matching and Sluicing when P-Stranding is not taken into account. On the other hand, however, even though P-stranding should be completely unacceptable overtly, regardless of Case-matching, under sluicing it is predicted to become as acceptable as P-pied-piping. As such, both P-stranding and P-pied-piping under sluicing should be equally acceptable and at ceiling level, but only when combined with case-matching, given all case-mismatching conditions - overt and under sluicing - should be unacceptable. This should result in a three-way interaction prediction between the factors of Case-matching, P-Stranding and Sluicing, such that the difference between P-stranding and P-pied-piping conditions under sluicing should be significantly modulated by the factor of Case-matching, whereas overtly, P-stranding and case-mismatching conditions should be equally unacceptable and at floor level, regardless of whether the two ungrammaticality effects are combined or not. Finally, given that case-mismatching is considered to be always unacceptable, regardless of sluicing, it is quite likely that sluicing conditions overall would not be significantly more highly rated than overt conditions, i.e. we would not predict a main effect of Sluicing.

With respect to an LF-Copying account, the predictions are the same as those of the PF-Deletion approach where PPs are assumed to be PF-islands. That is to say, case-matching P-stranding would be predicted to be as acceptable as P-pied-piping under sluicing and at ceiling level, thanks to there being no structure at the *e*-site and no specific phrasal category identity requirement assumed between correlate and remnant. LF-Copying's case-matching identity requirement, on the other hand, would predict case-mismatching to be always unacceptable and at floor level, regardless of the factor of P-Stranding, just like a PF-Deletion Hypothesis would. Hence, we would anticipate a main effect of Case-matching, with no two- or three-way interactions with P-stranding and/or Sluicing; an interaction of P-stranding with Sluicing; and no main effect of Sluicing.

Under a general pointer account, on the other hand, thanks to the eschewal of *e*-site structure and barring any other identity requirements between correlate and remnant, P-stranding and case-mismatching should both be completely unacceptable overtly, but acceptable under sluicing, i.e. one would predict a significant two-way interaction between the factors of Sluicing and P-Stranding, on the one hand, and Sluicing and Case-matching, on the other. This should also lead to a main effect of Sluicing, with sluicing conditions being overall significantly better than their overt counterparts. If a case-matching requirement between correlate and remnant is enforced, however, as under the pointer account of Nykiel (2013); Sag and Nykiel (2011), then a main effect of this factor would be anticipated, with case-mismatching always being deemed unacceptable, regardless of Sluicing and P-Stranding. If, furthermore, within the same school of thought, factors affecting general anaphora resolution, such as the complexity of correlate and remnant, also govern sluicing resolution, then one might anticipate that these would significantly interact with P-stranding acceptability under sluicing, such that more complex correlates are afforded significantly higher acceptability ratings when paired with P-stranding under sluicing compared to simpler ones. In other words, *which*-NPs would be predicted to have higher acceptability ratings for P-stranding under sluicing compared to *wh*-XPs, since the former refer to more complex correlates by default compared to the latter. With respect, finally, to an interaction between all three factors of Case-matching, P-Stranding and Sluicing, given that a) the pointer account does not make gradient acceptability predictions for anything other than the factor of remnant/correlate complexity, with conditions in general predicted to receive floor or ceiling level ratings overtly and under sluicing, and b) the factor of Case-matching is in no way predicted to be affected by either of the other factors, we would not anticipate a significant three-way interaction in any direction, whether under the more general pointer account or the more specific one of Nykiel (2013); Sag and Nykiel (2011).

From the point of view of a cross-linguistic generalisation, if we find that German shows an interaction between the factors of P-Stranding and Sluicing, as well as a main effect of P-Stranding, similarly to Greek, then this gives us a broader idea of what we must account for. Although it may also be the case that German presents with a completely different set of sluicing results to those of the original Greek experiment, with neither P-stranding nor case-mismatching being any more acceptable than their overt counterparts. If this is the case, then a cross-linguistic account of P-stranding being generally acceptable under sluicing would obviously be unrepresentative. Instead, it may be that Greek is - for some as-yet undefined set of reasons - an exceptional language, with its behaviour under sluicing being unrelated to the morpho-syntactic features it shares with German. If this proves to be true, then a number of further studies would be required to ascertain the exact reason for this cross-linguistic difference.

As in Experiment 1, a table representing the general theoretical predictions of each approach is presented below (Table 2.21).

Table 2.21: Experiment 4 Predictions

	Prediction of Significance	PF-Deletion	PF-Deletion (PP:PF-island)	LF-Copying	Pointer	Pointer (Nykiel)
1	Sluicing	✗	✗	✗	✓	✗
2	Case-Matching	✓	✓	✓	✗	✓
3	P-Stranding	✓	✗	✗	✗	✗
4	Sluicing*Case-Matching	✗	✗	✗	✓	✗
5	Sluicing*P-Stranding	✗	✓	✓	✓	✓
6	Sluicing*P-Stranding*Case-Matching	✗	✓	✗	✗	✗
7	Sluicing*P-Stranding*Complexity	✗	✗	✗	✗	✓

2.5.2 Methods

2.5.2.1 Experimental Design

As mentioned above and as in Experiment 1, the three binary factors of Sluicing (±), P-stranding (±) and Case-matching (±) were crossed to create 8 experimental conditions, presented in Table 2.22. Out of these conditions, the two unequivocally acceptable ones are the overt and covert case-matching, P-pied-piping ones (1 and 3). All conditions began with the structure in (72) and continued as shown in the table.

2.5.2.2 Method

As for all the previous Greek studies, a web-based acceptability judgement task was conducted, hosted on the Ibex Farm platform (www.spellout.net/ibexfarm). This time, participants were recruited via Prolific Academic®, a website specialising in connecting academics and other researchers with potential subjects for web-based studies. The website

Table 2.22: Experiment 4 Conditions and Examples

(72) Die erfolgreichen Schatzsucher wollen als nächstes
*Det.*NOM.PL. *successful.*NOM.PL. *treasure.seekers would.like.*3PL *for.Det.*N.ACC *next.*N.ACC
nach einem vergrabenen Goldschatz suchen und sie haben
*to some.*M.ACC *burried.*M.DAT *gold.treasure.*M.DAT *to.search and they have.*3PL
sich auch schon darüber geeinigt…
themselves PRT *already about.that reached.an.agreement*
'The successful treasure seekers would like to search for a buried treasure next, but they have not yet decided . . . '

	Condition	Example
1	Case-Matching, Non-Sluicing, P-Pied-Piping	. . . nach welchem sie suchen wollen. *to which.*M.DAT *they to.search would.like.*3PL '. . . for which (one) they would like to search.'
2	Case-Matching, Non-Sluicing, P-Stranding	. . . welchem sie suchen wollen nach. *which.*M.DAT *they to.search would.like.*3PL *to* '. . . which (one) they would like to search for.'
3	Case-Matching, Sluicing, P-Pied-Piping	. . . nach welchem. *to which.*M.DAT '. . . for which (one).'
4	Case-Matching, Sluicing, P-Stranding	. . . welchem. *which.*M.DAT '. . . which (one).'
5	Case-Mismatching, Non-Sluicing, P-Pied-Piping	. . . nach welcher sie suchen wollen. *to which.*M.NOM *they to.search would.like.*3PL '. . . for which (one) they would like to search.'
6	Case-Mismatching, Non-Sluicing, P-Stranding	. . . welcher sie suchen wollen nach. *which.*M.NOM *they to.search would.like.*3PL *to* '. . . which (one) they would like to search for.'
7	Case-Mismatching, Sluicing, P-Pied-Piping	. . . nach welcher. *to which.*M.NOM '. . . for which (one).'
8	Case-Mismatching, Sluicing, P-Stranding	. . . welcher. *which.*M.NOM '. . . which (one).'

would redirect participants to the Ibex platform and upon completion of the study they would be redirected back to Prolific to receive compensation (£8 per hour). On Ibex, the procedure was exactly the same as that in the Greek studies, i.e. participants would see a full sentence appear in a single line on-screen and were asked to rate it based on how acceptable they found it on a scale of 1 - 7, with 1 being completely unacceptable and 7 being completely acceptable. Although every effort was made to make all stimuli as plausible as possible, due to time limitations a separate plausibility study to judge the stimuli, as in the original Greek study, was not run. Participants were thus asked to not focus on the plausibility of the sentence's content, but rather on its form. Detailed instructions at the beginning of the study explained the scale and gave some example sentences. After 3 practice sentences, participants saw the main stimuli[20]. Before the start of the experiment, participants were presented with a digital copy of a study information sheet and consented to taking part knowing that the study posed no foreseeable risk to them and that they were able to withdraw at any time without repercussions[21].

To make sure that participants would not become otherwise distracted while reading the sentence, each stimulus would appear for a maximum of 45 seconds, after which it would automatically time out, with no response recorded. To further make sure that participants were paying adequate attention and comprehending the stimuli, comprehension questions appeared after 2/3 of experimental stimuli (only for the unequivocally acceptable conditions 1 and 3) and half of the grammatical fillers.

2.5.2.3 Items

2.5.2.3.1 Experimental Stimuli

Stimuli were created roughly following the basic pattern delineated in (73), for instance (74), a pattern similar to that used in the Greek experiment (word order not-withstanding).

(73) Main Clause Subject + Main Verb/Auxiliary + Preposition + Indefinite Internal Argument (+ Infinitive/Past Participle), Complementizer (+ Second Clause Subject) + Embedding Verb (+ NEGATION) (+ Infinitive/Past Participle) + Interrogative Pronoun (Dative).

(74) Die Maria hat mit jemandem getanzt, und warte bis du
*Det.*F.NOM *Maria.*F.NOM *has with someone.*M.DAT *danced and wait.*IMP *until you*
hörst wem.
*hear.*2SG *who.*M.DAT
'Maria danced with someone and just wait until you find out who.'

[20]Neither examples nor practice sentences included P-stranding or case-mismatching.

[21]In compliance with the EU GDPR 2018, no personally identifiable information was collected from participants other than the region within which they grew up.

This was a rough pattern, as the part following the Complementizer was obviously adapted for each condition. Further, non-discourse-related material, such as adverb phrases, were also used to embellish the main clause in order to create more natural-sounding sentences. Aside from this pattern, the stimuli were subject to a strict set of rules, similarly to the Greek study, in order to control for various potentially confounding factors which have been previously suggested to affect participants' judgements when it comes to sluicing.

Firstly, with respect to the assigned case, although in Greek the only case that can be assigned to a P's complement is Accusative, in German it can be either Accusative or Dative, depending on the verb plus P combination. In order to not introduce variation between items, we used only one case for the German replication and decided on Dative instead of Accusative simply due to the higher number of verb plus P options that it offered. Secondly, as in the Greek study, the main verb was chosen to only c-select for PP complements and not bare Dative-marked complements. This was particularly important in order to ensure that the bare *wh*-remnant could not possibly be interpreted as an alternative, direct object of the main verb, thereby potentially making it more acceptable than it would otherwise be judged as. Thirdly, with respect to feature-marking of the correlate and remnant, in the original Greek study we used only masculine NPs as internal arguments due to their overt case-marking suffixes clearly distinguishing between Nominative and Accusative in both numbers. Although German case-marking morphology can clearly distinguish between Dative and Nominative for all three genders, we decided to only use masculine NPs in this study as well in order to minimise variability between the two experiments. Overall, in German, as in Greek, the same suffix morpheme amalgamates the features of gender, case and number. For the case-mismatching conditions, as in the Greek study, these always featured Nominative instead of the grammatical Dative, as this is the case copular pivots must appear in, thus indicating whether a copular source could potentially be implied. All correlates and remnants were also presented in singular case. With respect to the gender of the main clause subject, we used only feminine- and neuter-marked arguments, in order to avoid potential competition interference in attempting to find the *wh*-XP's correlate in the antecedent based on shared morphosyntactic features (Caramazza et al., 1977; Criss & McClelland, 2006; Ehrlich, 1980; Hofmeister et al., 2013; Nairne, 1990, 2001, 2006; Oberauer & Lewandowsky, 2008). Moving on to the form of the remnant, as in Greek, in order to check for complexity effects (Nykiel, 2013), half of the *wh*-XPs were simple interrogative pronouns (*who*) and half were more complex *which*-NPs; of the latter, half contained an overtly expressed NP and half covert. With respect to the Ps themselves, equal numbers of four of the most common Ps were used, namely 'bei' (*near/for/at*), 'mit' (*with*), 'nach' (*after/to*), 'von' (*from/of*). Finally, an additional factor included here which was not examined in the original, Greek study was that of complementizer form. Specifically, to check whether different complementizers may encourage more or less parallelism (Lemke, Schäfer, & Reich, 2018), thereby possibly affecting how acceptable P-stranding or Case-mismatching may be, half of the items introduced the second clause with 'aber' (*but*) and half with 'und' (*and*). All of the above factors were controlled for and equally distributed across items to yield the table shown in (2.23).

In addition to these constraints, the stimuli were, furthermore, controlled for sentence length as much as possible. Given the nature of our manipulation, there was considerable variation between elliptical and non-elliptical conditions. The elliptical conditions had a Mean of 20.25 (SD: 2.6); the non-elliptical conditions had a Mean of 25.25 (SD: 3.1); with the fillers having a Mean of 20 and a considerably larger degree of variation to cover both types of conditions (SD: 6.079) and not be significantly different from them.

2.5.2.4 Thermometer Items

An additional set of items which we decided to include in this study beyond our regular experimental and filler items were what can be termed *thermometer items*. These are based on an interesting notion first proposed in Featherston (2009): the use of an open-ended scale for acceptability judgements which would contain certain anchor points along the scale. These anchor points would be specific linguistic examples of varying acceptability which would serve to help 'ground' subjects and to give them specific points of comparison when judging subsequent items, the actual experimental items. This could be considered similar in some ways to the Magnitude Estimation method, but without claiming the use of ratio judgements or allowing different stimuli to be used as measures of comparison for each subject. Featherston (2009) gives 5 pairs of sentences in German which could serve as these 'cardinal well-formedness values' and attempts to show how these aid us in understanding how participants are actually rating the experimental stimuli we are interested in. In a conference talk in 2018, Gerbrich & Schreier similarly present 5 such stimuli for use with English.

Neither study explains how they created these five stimuli types, i.e. what forms of grammatical degradation were used or if they could be explained in terms of quantification (e.g. one stimulus has an instance of case mismatching, another one has case-mismatching plus unacceptable word order etc.). From an examination of both sets of stimuli, they appear to be quite randomly generated. Although using such thermometer items is an intriguing concept, and one that would be very interesting to look into in more detail, it is also unfortunately quite unclear how it would work in practice, as neither study proposes how to use these items other than to simply insert all of them amongst the filler items and use their responses to compare all experimental items against. This in and of itself is inherently problematic. As the study progresses, participants see more and more items of varying acceptability and are very likely to change their inner 'anchor' points for what

Table 2.23: Experiment 4 Stimuli Organisation

<table>
<tr><th>Preposition</th><th colspan="2">Remnant Form</th><th>Complementizer</th><th colspan="5">Total</th></tr>
<tr><td rowspan="6">bei</td><td colspan="2" rowspan="2">who</td><td>und</td><td colspan="2">4</td><td rowspan="2">8</td><td rowspan="6">16</td><td rowspan="24">64</td></tr>
<tr><td>aber</td><td colspan="2">4</td></tr>
<tr><td rowspan="4">which</td><td rowspan="2">which</td><td>und</td><td>2</td><td rowspan="2">4</td><td rowspan="4">8</td></tr>
<tr><td>aber</td><td>2</td></tr>
<tr><td rowspan="2">which-NP</td><td>und</td><td>2</td><td rowspan="2">4</td></tr>
<tr><td>aber</td><td>2</td></tr>
<tr><td rowspan="6">mit</td><td colspan="2" rowspan="2">who</td><td>und</td><td colspan="2">4</td><td rowspan="2">8</td><td rowspan="6">16</td></tr>
<tr><td>aber</td><td colspan="2">4</td></tr>
<tr><td rowspan="4">which</td><td rowspan="2">which</td><td>und</td><td>2</td><td rowspan="2">4</td><td rowspan="4">8</td></tr>
<tr><td>aber</td><td>2</td></tr>
<tr><td rowspan="2">which-NP</td><td>und</td><td>2</td><td rowspan="2">4</td></tr>
<tr><td>aber</td><td>2</td></tr>
<tr><td rowspan="6">nach</td><td colspan="2" rowspan="2">who</td><td>und</td><td colspan="2">4</td><td rowspan="2">8</td><td rowspan="6">16</td></tr>
<tr><td>aber</td><td colspan="2">4</td></tr>
<tr><td rowspan="4">which</td><td rowspan="2">which</td><td>und</td><td>2</td><td rowspan="2">4</td><td rowspan="4">8</td></tr>
<tr><td>aber</td><td>2</td></tr>
<tr><td rowspan="2">which-NP</td><td>und</td><td>2</td><td rowspan="2">4</td></tr>
<tr><td>aber</td><td>2</td></tr>
<tr><td rowspan="6">von</td><td colspan="2" rowspan="2">who</td><td>und</td><td colspan="2">4</td><td rowspan="2">8</td><td rowspan="6">16</td></tr>
<tr><td>aber</td><td colspan="2">4</td></tr>
<tr><td rowspan="4">which</td><td rowspan="2">which</td><td>und</td><td>2</td><td rowspan="2">4</td><td rowspan="4">8</td></tr>
<tr><td>aber</td><td>2</td></tr>
<tr><td rowspan="2">which-NP</td><td>und</td><td>2</td><td rowspan="2">4</td></tr>
<tr><td>aber</td><td>2</td></tr>
</table>

constitute better or worse stimuli within the specific study setting. This is, after all, one of the reasons for order effects in experimental results, as participants become more and more used to the experimental manipulations over time. In this sense, it is one thing to talk about boiling and freezing points on a temperature scale[22], which are easily observed and independently attestable, and quite another to try and find one, two or five such reference points in linguistics, which would not only be universally equally judged, but also readily available to participants at any time. In order for us to use certain stimuli as reference points, we would have to either have these sentences constantly present on-screen at the same time as the other, experimental sentences to be judged, something which would make for quite a confusing, visually noisy and time-consuming process; or we would have to drill participants into giving these 5 sentences specific judgements at the beginning of the study.

Bearing in mind these objections, we have included in our fillers the 5 pairs of German sentences provided in Featherston (2009) (included in Appendix), both as potential reference points, but also to see if participants do indeed give them uniform judgements.

2.5.3 Results

2.5.3.1 Experimental Items

This study, along with all analyses presented here, was pre-registered and frozen prior to data collection on the Open Science Framework website and can be viewed at `https://osf.io/58b6c/`.

The data of 95 participants were collected. Participants were aged 18 and above, right-handed, native German speakers, with German as their first language and no history of learning disabilities, neurological or psychiatric disorders. All these criteria were self-attested and used to filter out non-eligible participants by default on Prolific Academic. We also collected data on which particular region of Germany or Austria participants grew up in, to check for potential dialectal variability as had been found in Greek. Prior to analysis, the data were cleaned, excluding those of participants who responded with less than 80% accuracy to all comprehension questions. In this manner, 14 participants (i.e. 14.7% of the total number) were excluded from the final analysis, leaving a total of 81 subjects[23]. All responses made under less than 100ms (8 in total or 0.05% of the total number of data points) were also excluded as erroneous. Outliers were calculated per condition per participant at $2.5 * SDs$ from the Mean and Windsorised (i.e. replaced by the minimum or maximum allowable

[22] A metaphor used by Featherston (2009).

[23] There is quite a large difference in participant exclusion between this study and the original Greek one (i.e. 14 vs. 41). We believe this is probably due to the different method utilised in recruiting participants, i.e. in one instance they were simply acquaintances/friends who did not necessarily complete the entire study in one go or may have been distracted whilst responding. Another incentive to complete the German study with a larger degree of attention was that all participants were paid a set fee for their participation, as opposed to the participants for the Greek study who were entered into a lottery draw for a single prize.

value). After practice items and fillers were excluded, this left a total of 5,172 data points or approximately 646 data points per condition.

A maximal linear mixed effects regression model was fitted to the data (as per Barr et al. (2013)), including the factors of P-stranding (±), Sluicing (±) and Case-matching (±) as fixed effects and subjects and items as random effects (with random slopes and intercepts assumed for each). The same model was fitted with raw scores, log-transformed scores, z-scores and log-transformed z-scores as dependent variables. Each time, the same main effects and interactions were found, with some small differences in the size of the effects, but no significant difference in model fit. For this reason simply the results of the model with raw scores as DV are presented in-text here, with all others shown in Table 2.25. A summary table and barplots depicting acceptability Means per condition with SEM are given in Table 2.24 and Figure 2.8, respectively, with Figure 2.9 showing the average filler results for comparison purposes. The model showed significant two- and three-way interactions, as well as main effects for all three fixed factors, described below, with post-hoc comparisons (Tukey-adjusted) for significant two- and three-way interactions carried out by comparing estimated marginal means (EMMs) via the "emmeans" R package.

With respect to the effects of each of these factors separately, a main effect was found for a) P-Stranding ($t > 19.21; p < 1e{-}16$), with P-pied-piping conditions (Mean = 4.94; SEM = 0.04) being overall significantly more acceptable than their P-stranding counterparts (Mean = 3.40; SEM = 0.04; Mean difference = 1.54); b) Case-matching ($t > 17.191; p < 1e - 16$), with case-matching conditions (Mean = 5.05; SEM = 0.04) being overall significantly more acceptable than case-mismatching ones (Mean = 3.29; SEM = 0.04; Mean difference = 1.76); and c) Sluicing ($t > 10.04; p < 1e - 16$), with sluicing conditions (Mean = 4.50; SEM = 0.04) being overall significantly more acceptable than non-sluicing ones (Mean = 3.83; SEM = 0.05; Mean difference = 0.67). Moving on to the two-way interactions, the model showed a significant two-way interaction between P-Stranding and Sluicing ($t > 9.93, p < 1e - 16$), with post-hoc comparisons showing that a) the difference between P-pied-piping and P-stranding overtly (Mean difference: 2.22) was significantly larger than the equivalent difference under sluicing (Mean difference: 0.89$p < .001$) and b) there was a significant difference between non-sluicing and sluicing conditions when paired with P-stranding (Mean difference: 1.32; $p < .0001$), but not when paired with P-pied-piping (Mean difference: $0.01; p > .856, n.s.$). Together, these results indicate that - when Case-matching is not taken into account - although P-pied-piping is still significantly better than P-stranding both overtly and under sluicing, sluicing makes P-stranding significantly better than it is overtly, with no change in acceptability results for P-pied-piping. On the other hand, there was no significant two-way interaction between Case-matching and Sluicing ($t < 1.44, p > .149, n.s.$), indicating that - when the factor of P-Stranding is not taken into account - the size of the difference between case-matching and mismatching conditions is not significantly affected by the factor of Sluicing, or in other words the context of sluicing does not appear to ameliorate case-mismatching. Moving forward, a significant two-way interaction was found between Case-matching and P-Stranding ($t > 16.92; p < 1e - 16$), with post-hoc comparisons showing that, when the factor of Sluicing was averaged over, a) the difference between case-matching and mismatching was significantly larger for P-pied-piping conditions (Mean difference $= 3.30$) than for P-stranding ones (Mean difference $= 0.22, p < .0001$); and b) the difference between P-pied-piping and P-stranding conditions was significantly greater with case-matching (Mean difference $= 3.09$) than with case-mismatching (Mean difference $= 0.01, p < .0001$). In other words, case-mismatching appeared to have a greater effect on P-pied-piping conditions than on P-stranding ones.

Finally, a significant three-way interaction was found between all three factors ($t > 3.64, p < .0002$). Post-hoc comparisons allowed this complex interaction to be teased apart. When we focus on each level of Case-matching separately, although for case-matching conditions the difference between P-pied-piping and P-stranding was significant both overtly (Mean difference overtly: $3.58; p < .0001$) and under sluicing (Mean difference under sluicing: $2.59; p < .0001$), the size of this difference was significantly affected by the factor of Sluicing, with the difference being greater overtly than under sluicing (Mean size difference: $0.98; p < .0001$). For case-mismatching conditions, on the other hand, the difference between P-pied-piping and P-stranding was again significant both overtly (Mean difference: $0.86; p < .0001$) and under sluicing (Mean difference: $0.82, p < .0001$), however the factor of Sluicing did not significantly affect the size of this difference (Mean size difference: $0.041; p > .09 n.s.$), but rather its direction (Mean size difference when taking into consideration direction of difference $1.67; p < .0001$), with case-mismatching P-stranding being significantly *less* acceptable than P-pied-piping overtly, but significantly *more* acceptable than P-pied-piping under sluicing. In fact, case-mismatching was the least acceptable when paired with P-stranding overtly and the most acceptable when paired with P-stranding under sluicing. When we focus on each level of P-Stranding separately, for P-pied-piping conditions there is a significant difference between case-matching and case-mismatching both overtly (Mean difference: $3.20; p < .0001$) and under sluicing (Mean difference: 3.38), however the factor of Sluicing did not significantly affect the size of this difference (Mean size difference: $0.18; p > .071, n.s.$) with the two being relatively close to each other. For P-stranding conditions, although there is a significant difference between case-matching and case-mismatching overtly (Mean difference: $0.48; p < .0001$), there is almost no difference whatsoever between case-matching and case-mismatching under sluicing (Mean difference: $0.01; p > .999, n.s.$), with the factor of Sluicing thus having a significant effect on the size of the difference between case-matching and case-mismatching (Mean size difference: $0.53; p < .0001$). Together with the previous results, this would indicate that P-stranding and case-mismatching together have an additive effect on acceptability ratings overtly, but not under sluicing. Let us now examine the difference between overt and sluicing conditions for P-pied-piping and P-stranding. Specifically, when we focus on P-pied-piping conditions only, there is no significant difference between overt and sluicing conditions when paired with either case-matching (Mean difference: $0.08; p > .809, n.s.$) or case-mismatching (Mean difference: $0.07; p > .869, n.s.$), with the factor of Case-matching not having a significant effect on the size of this difference either (Mean size difference: $0.01; p > .999, n.s.$). On the other hand, if we focus on P-stranding conditions only, there is a significant difference between overt and sluicing conditions, both when paired with case-matching (Mean difference: $1.07; p < .0001$) and case-mismatching (Mean difference: $1.60; p < .0001$), with the size of this difference being significantly greater with case-matching than case-mismatching (Mean size difference: $0.53; p < .0001$). Moving on, let us examine the difference between overt and sluicing conditions for case-matching and case-mismatching. Specifically, when we focus on case-matching conditions only, the difference between overt and sluicing conditions is not significant when paired with P-pied-piping (Mean difference: $0.08; p > .809, n.s.$), however it is significant when paired with P-stranding (Mean difference: $1.07; p < .0001$), with the size of this difference being significantly greater with P-stranding than P-pied-piping. When we focus on case-mismatching conditions only, we see the same pattern of effects, with the difference between overt and sluicing conditions being non-significant when paired with P-pied-piping (Mean difference: $0.07; p > .869, n.s.$), but significant when paired with P-stranding (Mean difference: $1.60; p < .0001$), with the size of this difference being significantly greater with P-stranding than P-pied-piping (Mean size difference: $1.67; p < .0001$). These results, together with the significant two-way interaction we saw above between P-Stranding and Sluicing, lend further evidence to the hypothesis that P-stranding becomes significantly better under sluicing than it is overtly, with no such amelioration for P-pied-piping. Finally, and repeating some of the already examined differences for symmetry, examining each level of Sluicing separately, for overt conditions, there was a significant difference between both case-matching P-pied-piping and P-stranding conditions (Mean difference: $3.57; p < .0001$), on the one hand, and case-mismatching P-pied-piping and P-stranding conditions (Mean difference: $0.86; p < .0001$), on the other hand, the size of this difference was significantly affected by the factor of Case-matching (Mean size difference: $2.74; p < .0001$), with the difference between P-pied-piping and P-stranding being significantly greater for case-matching conditions than case-mismatching ones. When we examine sluicing conditions, this pattern is quite similar; once again we find a significant difference between P-pied-piping and P-stranding when paired with case-matching (Mean difference: $2.59; p < .0001$), and when paired with case-mismatching (Mean difference: $0.82, p < .0001$), with the size of the difference between P-pied-piping and P-stranding being significantly affected by the factor of Case-matching (Mean size difference: $1.77; p < .0001$).

With respect to the additional factors included in the study design, i.e. Complementizer

Figure 2.8: Experiment 4: Acceptability Ratings Barplot

Explanatory Notes: Mean response per condition with SEM error bars

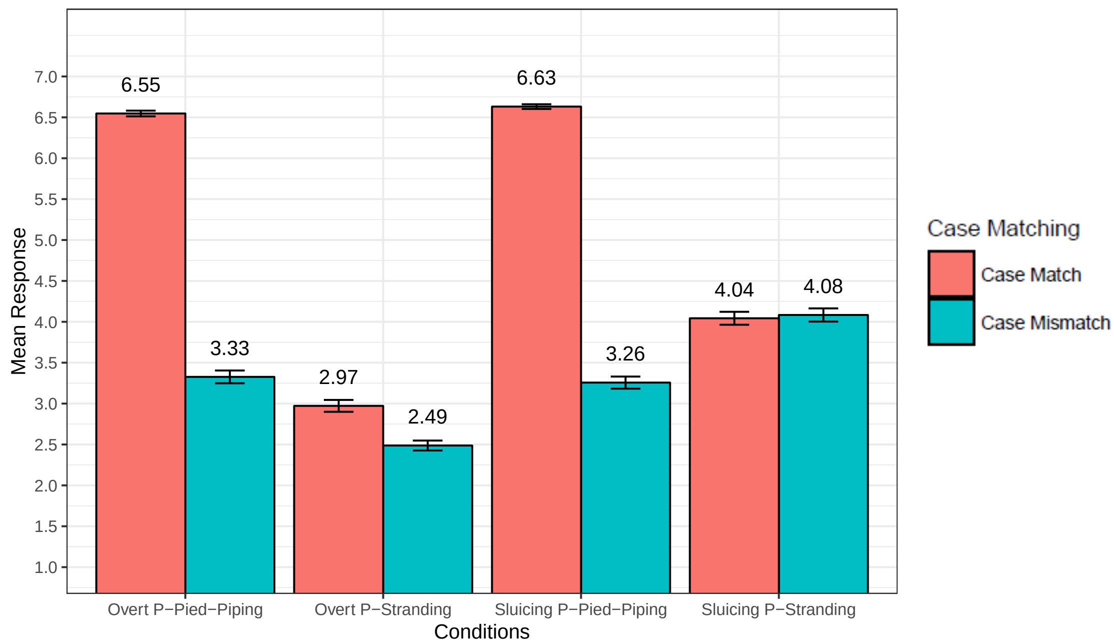


Table 2.24: Experiment 4: Acceptability Ratings Numerical Summary

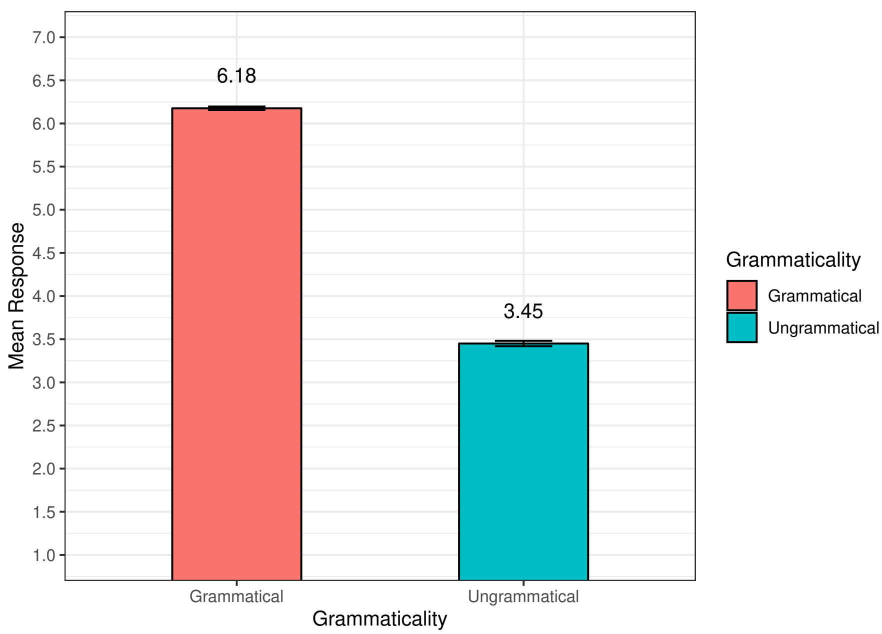
	Condition	Response	SD	SEM	95% CI
1	Case-Matching Non-Sluicing P-Pied-Piping	6.547	0.883	0.035	0.068
2	Case-Matching Non-Sluicing P-Stranding	2.972	1.845	0.073	0.143
3	Case-Matching Sluicing P-Pied-Piping	6.631	0.732	0.029	0.057
4	Case-Matching Sluicing P-Stranding	4.043	2.014	0.079	0.155
5	Case-Mismatching Non-Sluicing P-Pied-Piping	3.327	1.980	0.078	0.153
6	Case-Mismatching Non-Sluicing P-Stranding	2.487	1.544	0.061	0.119
7	Case-Mismatching Sluicing P-Pied-Piping	3.257	1.879	0.074	0.145
8	Case-Mismatching Sluicing P-Stranding	4.083	2.049	0.081	0.158

Table 2.25: Experiment 4: Main Effects & Interactions

		Raw Scores		z-Scores		$Log10$-Scores	
	Factor	t-value	p-value	t-value	p-value	t-value	p-value
1	Case-Matching	17.190	< 0.0001	21.946	< 0.0001	13.010	< 0.0001
2	Sluicing	10.043	< 0.0001	10.139	< 0.0001	10.098	< 0.0001
3	P-Stranding	19.205	< 0.0001	22.569	< 0.0001	15.239	< 0.0001
4	Case-Matching*Sluicing	1.441	> 0.149	1.362	> 0.173	1.685	$> .092$
5	Case-Matching*P-Stranding	16.924	< 0.0001	22.510	< 0.0001	12.560	< 0.0001
6	P-Stranding*Sluicing	9.933	< 0.0001	10.232	< 0.0001	10.026	< 0.0001
7	Case-Matching*Sluicing*P-Stranding	3.646	< 0.0003	3.631	< 0.0003	3.357	< 0.0008

Figure 2.9: Experiment 4: Fillers Mean Acceptability Ratings Barplot

Mean response per condition with SEM error bars.



Form ('and' vs. 'but'), Remnant Type (*who* vs. *which*-NP, and of the latter, *which*-NP vs. *which*) and P Type (*bei*, *mit*, *nach*, *von*) none showed a significant main effect nor interaction with any other factor (all t's < 1.478; all p's $> .139, n.s.$). This would indicate that none of these factors appear to be having any effect on P-stranding or case-mismatching results under sluicing. For the *which*-NP remnant type in particular, with respect to the repeated or omitted NP, this lack of difference in results appears contrary to what the 'repeated name penalty' would predict.

Finally, following Nakagawa and Schielzeth (2013), we can assess the model's fit by calculating its conditional R^2 to be 0.83, meaning that this model including the three fixed effects factors and two random effects factors, can account for over 83% of the variation in the data, i.e. a large effect size, following Sullivan and Feinn (2012). A power analysis (following Westfall et al. (2014)) estimated the study power at 1, meaning the same result would be obtained 100% of the time, a very robust result.

2.5.3.2 Thermometer Item Results

For comparison purposes, the results of the 'Thermometer Items' from Featherston (2009) are reported below. In the original paper, no absolute ratings were given per item, however the two items in each group were reported to be given overall uniform judgements from participants in each experiment they were included in. Group A was consistently judged as the most acceptable, with this acceptability rating falling through Groups B to E, with Group E being judged as the least acceptable.

Overall, the results for these items in our experiment also appear to follow the trend reported in Featherston (2009), with judgements being the highest for Group A and the lowest for Group E (Figure 2.10). That having been said, for Groups C, D and E, there does appear to be substantial amount of variability in participants' responses, as opposed to the more uniform judgement predicted in the aforementioned paper, as evident from the large SD for each (Table 2.26). A scatterplot is also provided in Figure 2.11 demonstrating how these responses per Group varied per subject (jitter added so that each subject's response can be seen more clearly). Given this large variability, I do not believe that a direct comparison of all five groups against the results of the experimental stimuli would be particularly enlightening as to the meaning of the latter. As such, I have simply included these results for the reader to compare them to those of the main stimuli, if desired. I will simply observe here that numerically the two unequivocally 'acceptable' conditions in our experiment (namely Case-matching with P-pied-piping, overtly and under sluicing) appear closest to the second most acceptable Thermometer Group, i.e. 'B', whereas the least acceptable experimental condition (Case-mismatching, Non-Sluicing, P-stranding) appears closest to the second least acceptable Thermometer Group, i.e. 'D'.

Table 2.26: Thermometer Items: Acceptability Results Summary

	Thermometer Item	Response	SD	SEM	95%CI
1	Group A	6.951	0.366	0.029	0.057
2	Group B	6.469	1.138	0.089	0.177
3	Group C	3.821	2.106	0.165	0.327
4	Group D	2.358	1.598	0.126	0.248
5	Group E	1.809	1.455	0.114	0.226

Figure 2.10: Thermometer Items: Acceptability Ratings Barplots

Explanatory Notes: Mean response per condition with SEM error bars

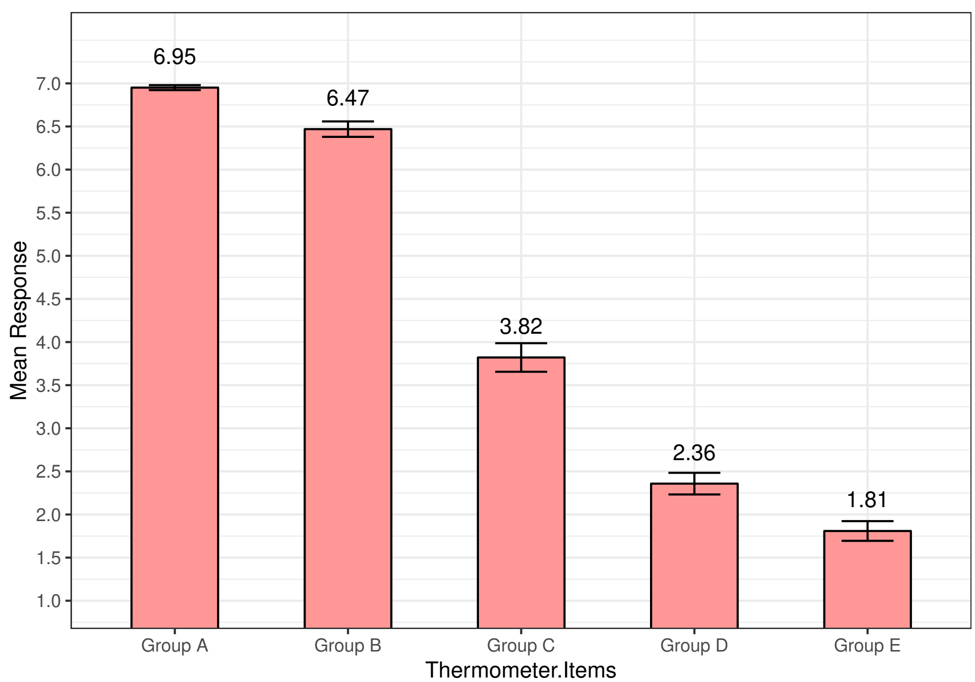
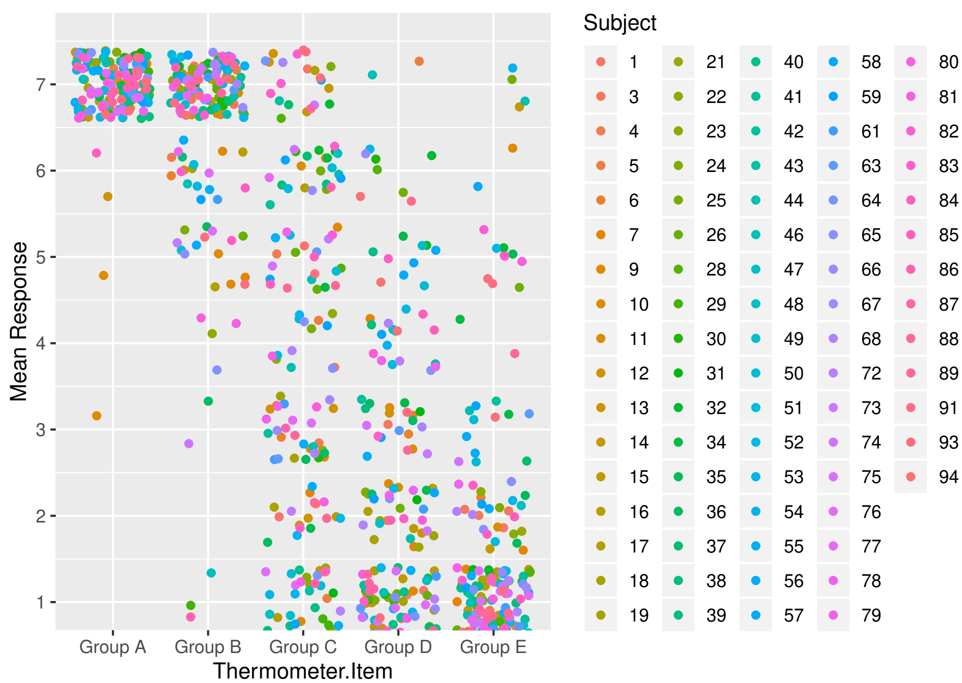


Figure 2.11: Thermometer Items: Acceptability Ratings Scatterplot per Subject

Explanatory Notes: Mean response per subject with added jitter



2.5.3.3 Comparison of Greek and German

In order to get a better idea of the differences in effects between Greek and German, a comparative analysis of the two datasets was also run. Naturally, this comparison is not ideal, given we are comparing the results of two separate studies using different subjects and items. That having been said, fitting a linear mixed effects model on the combined datasets of both languages, with 'Language' as a between-subjects factor in an additional exploratory analysis can still give us some idea of the differences between the two sets of results.

A linear mixed effects model was thus fitted to the dependent variable of raw responses, with the four fixed effects factors of Sluicing (±, Sluicing vs. Non-Sluicing continuation), P-stranding (±, P-stranding vs. P-pied-piping), Case-matching (±, Case-matching vs. Case-mismatching) and Language (Greek vs. German), with Sluicing, P-stranding and Case-matching as within-subjects factors and Language as a between-subjects factor. The random effects of subjects and items were also added to the model, with random slopes and intercepts assumed for the three within-subjects fixed effects factors (i.e. Language was not considered to have different effects for each subject and item). Given the results of this model can become complicated, the reader is referred to tables 2.27 and 2.28 for full details of significance[24], as well as relevant Means and SEM for each language, with a useful barplots graph comparing the two languages in Figure 2.12. The model converged finding the same main effects and interactions for Sluicing, P-stranding and Case-matching as found in each study and magnified. Given the same effects were found in each study and were, furthermore, in the same direction, these results were to be expected here and are reported in Table 2.28. More interestingly, however, the model also found a significant main effect of Language, with German on average receiving higher acceptability ratings than Greek (German Mean: 4.17; SEM: 0.031; Greek Mean: 3.95; SEM: 0.03; Mean difference: 0.22). This main effect was further modulated by significant two-way (Case-matching * Language; Sluicing * Language), three-way (Case-matching * Sluicing * Language; Case-matching * P-stranding * Language; P-stranding * Sluicing * Language) and four-way interactions (Case-matching * Sluicing * P-stranding * Language).

Specifically, the model showed a significant two-way interaction between Case-matching and Language ($t > 9.04; p < 1e-16$) with post-hoc comparisons showing that the difference between Greek and German was significant for case-mismatching conditions (Mean difference: $0.44; p < .0001$), but not for case-matching ones (Mean difference: $0.02; p > .739, n.s.$), with Language thus having a significant effect on the size of this difference (Mean size difference: $0.46, p < .0001$). Furthermore, the difference in acceptability between case-matching and case-mismatching conditions was significant for both Greek (Mean difference:

[24]As in all other analyses in this chapter, the same LME model was applied to raw scores, $z-$score- and $Log10-$score transformations. There was almost no difference between these models, other than for Case-Matching*P-Stranding*Language, which was not significant with $Log10-$scores.

$2.23; p < .0001$) and German (Mean difference: $1.76; p < .0001$), but the size of this difference was significantly greater for Greek than German (Mean size difference: $0.46; p < .0001$). The model also showed a significant two-way interaction between Sluicing and Language ($t > 6.98; p < 3.04e - 12$), with the difference between overt and sluicing conditions being significant for both Greek (Mean difference: $1.05; p < .0001$) and German (Mean difference: $0.67; p < .0001$) and the size of this difference being significantly bigger for Greek compared to German (Mean size difference: $0.38; p < .0001$). Furthermore, the difference in acceptability ratings between the two languages was significant for both overt (Mean difference: $0.519; p < .0001$) and sluicing conditions (Mean difference: $0.148; p < .0005$) with the size of this difference being significantly bigger for overt than sluicing conditions (Mean size difference: $0.38; p < .0001$). Interestingly, there was no significant two-way interaction between the factors of P-Stranding and Language ($t < 1.75; p > .08, n.s.$), indicating that as far as P-Stranding was concerned, the two languages had overall similar results.

Moving on to the three-way interactions, as mentioned above, the model found significant three-way interactions between all four variables in all combinations. Given a detailed in-text analysis of each of these complex interactions would be not only tiring but also potentially confusing for the reader, detailed contrasts are presented instead in Tables B.1, B.2, B.3 and B.4 in the Appendix for the interactions of a) Case-matching, P-Stranding and Sluicing, b) Case-matching, Sluicing and Language; c) Case-matching, P-Stranding and Language; and d) P-Stranding, Sluicing and Language, respectively, with only the most interesting contrasts described in-text. Specifically, beginning with the significant three-way interaction of Case-matching, P-Stranding and Sluicing ($t > 4.55; p < .0001$), post-hoc comparisons showed that when we collapse across the two languages, although all other comparisons were significant, the difference between sluicing and overt conditions when paired with P-pied-piping was not significant for either case-matching (Mean difference: $0.16; p > .335, n.s.$) or case-mismatching conditions (Mean difference: $0.01, p > .999, n.s.$). This shows that across both languages (as also seen previously in the results sections of each language separately), although P-stranding conditions become significantly better under sluicing, there was no difference in acceptability ratings between overt and sluicing structures for P-pied-piping conditions.

Moving on to the significant three-way interaction of Case-matching, Sluicing and Language ($t > 7.82; p < .0001$), when we focus only on overt conditions, the difference in acceptability results between Greek and German was significant for both case-matching conditions (Mean difference: $0.41; p < .0001$) and case-mismatching conditions (Mean difference: $0.39; p < .0001$), with the size of this difference not being significantly greater for either case-matching or case-mismatching (Mean size difference: $0.02; p > .839, n.s.$). On the other hand, when we focus on sluicing conditions, the difference in acceptability results between the two languages was again significant for both case-matching (Mean difference:

$0.44; p < .0001$) and case-mismatching conditions (Mean difference: $0.51; p < .0001$), however the size of this difference was significantly bigger for case-mismatching conditions than case-matching ones (Mean size difference: $0.96; p < .0001$). Along with the results we saw in both studies, these results confirm that under sluicing, although both case-matching and case-mismatching were significantly more highly rated in German than in Greek, this was particularly true for case-mismatching conditions.

When examining the significant three-way interaction of Case-matching, P-Stranding and Language now ($t > 5.73; p < .0001$), when we focus only on P-pied-piping conditions, the difference in acceptability ratings between Greek and German is not significant for case-matching conditions (Mean difference: $0.13; p > .245, n.s.$), however it is significant for case-mismatching ones (Mean difference: $0.61; p < .0001$), with the factor of Case-matching thus having a significant effect on the size of this difference (Mean size difference: $0.48; p < .0001$). When we focus only on P-stranding conditions, we see a similar significance pattern, with the difference between the two languages being significant for case-mismatching conditions, but not case-matching ones. Together with the previous three-way interaction seen, along with the lack of a significant two-way interaction between P-Stranding and Language seen above, this further elucidates the fact that both P-pied-piping and P-stranding conditions, at least when paired with case-matching, did not behave significantly differently between the two languages.

Moving on to the final three-way interaction between P-Stranding, Sluicing and Language ($t > 4.79; p < .0001$), when we focus on overt conditions only, the difference in acceptability ratings between Greek and German was significant both for P-pied-piping (Mean difference: $0.35; p < .0001$) and P-stranding (Mean difference: $0.46; p < .0001$), with the size of this difference not being significantly affected by the factor of P-Stranding (Mean size difference: $0.11; p > .276, n.s.$). On the other hand, for sluicing conditions, although the difference between the two languages was significant again for P-pied-piping conditions (Mean difference: $0.33; p < .0001$), it was not significant for P-stranding conditions ($0.04; p > .9999, n.s.$), with the factor of P-stranding thus having a significant effect on the size of this difference (Mean size difference: $0.29; p < .0001$). This result shows that P-stranding under sluicing in German is not rated significantly differently than in Greek when the factor of Case-matching is not taken into consideration.

Finally, let us look at the significant four-way interaction between all the fixed effects factors: Case-matching, P-Stranding, Sluicing and Language ($t > 10.62; p < .0001$). Leaving aside contrasts within each language which have already been delved into when discussing the results of Greek (2.1.3) and German (2.5.3.1) separately from each other, and referring the reader to a detailed four-way contrast table in the Appendix (Table B.5), we will instead focus again in-text only on the contrasts directly comparing the results of Greek and German. In analysing the significant three-way interactions above, we have seen evidence for a) case-matching and case-mismatching behaving significantly differently between the two languages when P-Stranding is not taken into consideration, with b) case-mismatching being more acceptable under sluicing for German compared to Greek; c) P-pied-piping and P-stranding behaving similarly across both languages when Sluicing is not taken into consideration; and d) when the factor of Sluicing *is* taken into consideration, but that of Case-matching is not, P-pied-piping conditions are significantly different overtly and under sluicing between the two languages, however P-stranding under sluicing appears to behave similarly. By analysing this complex four-way interaction now, we will be able to put together this complex mosaic of results into a clearer full-size picture, establishing what is actually occurring when all four factors are taken into consideration at the same time. Specifically, when we focus on overt P-pied-piping conditions, the difference in acceptability ratings between Greek and German is not significant when paired with case-matching (Mean difference: $0.15; p > .891, n.s.$), however it *is* significant when paired with case-mismatching (Mean difference: $0.67; p < .0001$), with the factor of Case-matching thus having a significant effect on this difference (Mean size difference: $0.52; p < .0001$). On the other hand, when we focus on overt P-stranding conditions, the difference between Greek and German ratings is significant both when paired with case-matching (Mean difference: $0.74; p < .0001$) and case-mismatching (Mean difference: $0.51; p < .0001$) with the size of this difference being significantly greater for case-matching than case-mismatching conditions (Mean size difference: $0.23; p < .01$). When we look at sluicing conditions, we see a similar pattern emerge. Specifically, when we focus only on sluicing conditions with P-pied-piping, the difference between Greek and German is again significant when paired with case-mismatching (Mean difference: $0.55; p < .0001$), but not with case-matching (Mean difference: $0.11; p > .996, n.s.$), with the factor of Case-matching thus having a significant impact on the size of this difference (Mean size difference:). On the other hand, when we focus only on sluicing conditions with P-stranding, the difference between Greek and German is significant when paired with both case-matching (Mean difference: $0.95, p < .0001$), with Greek case-matching conditions being rated significantly more highly than their German counterparts, and case-mismatching (Mean difference: $0.88; p < .0001$), with Greek case-mismatching conditions being rated significantly lower than their German counterparts, with no significant effect of Case-matching on the size of this difference.

Overall, this rather convoluted set of results can be summarised as follows: a) overall, German speakers rated almost all conditions more highly than Greek speakers, with the exception of P-stranding under sluicing; b) the behaviour of the factor of Case-matching overtly is similar across both languages, however under sluicing, case-mismatching conditions in particular are significantly more acceptable in German than in Greek; c) this case-mismatching difference between Greek and German is true both for P-pied-piping and P-stranding conditions, with a significantly larger difference between the two for P-stranding conditions; d) that having been said, when we ignore the factor of Case-matching or focus only on case-matching conditions, with respect to the factor of P-Stranding, the two languages do not behave significantly differently from each other; that is to say, e) in both Greek and German, P-pied-piping is acceptable both overtly and under sluicing, with no significant difference between overt and sluicing conditions; and f) in both Greek and German, P-stranding, on the other hand, is considered unacceptable overtly, however it becomes significantly more acceptable under sluicing; g) this difference, furthermore, between overt and sluicing P-stranding conditions is more pronounced in Greek, with both P-stranding overtly being rated lower and P-stranding under sluicing being rated higher in Greek than in German.

Finally, when we consider Sluicing, P-Stranding and Case-matching together, the grammaticality violations of case-mismatching together with P-stranding appear to have an additive effect overtly in both languages (numerically in Greek; significantly in German), although under sluicing they have a significant additive effect only in Greek, with both P-stranding and P-pied-piping with case-mismatching being rated identically in German. This rather perplexing difference between languages is definitely something worth looking into more with follow-up studies and is something we will return to in the next chapter when exploring these results in the context of a noisy channel hypothesis.

Following Nakagawa and Schielzeth (2013), conditional R^2 was found to be 0.62, meaning that this model including the four fixed effects factors and two random effects factors, can account for approximately 62% of the variation in the combined datasets, i.e. a large effect size, following Sullivan and Feinn (2012). A power analysis (following Westfall et al. (2014)) estimated the study power at 1, meaning the same result would be obtained 100% of the time, a very robust result.

2.5.4 Discussion

In summary, with respect to the German results, starting from the least surprising ones, we can see that case-mismatching in German is overall less acceptable than case-matching both under sluicing and overtly; P-pied-piping is also always more acceptable than P-stranding, regardless of sluicing; both of these results were unequivocally predicted by all theoretical approaches. That having been said, things become interesting when we look at the interaction of these two factors with sluicing. Specifically, considering first the case-matching conditions only, it is clear that in German, as in Greek, P-stranding becomes significantly more acceptable under sluicing than it is overtly, although still not being as acceptable as P-pied-piping. This result does not fit a purely syntactic approach to sluicing, but rather one that would eschew structure at the *e*-site or rely upon an interface of syntax with sentence processing to explain it. As such, a pointer account could capture this amelioration via sluicing, however it cannot explain why it is still significantly worse than P-pied-piping. The

Figure 2.12: Greek & German: Acceptability Ratings

Explanatory Notes: Mean response per condition with SEM error bars for Experiments 1 and 4; faceted by Case-matching

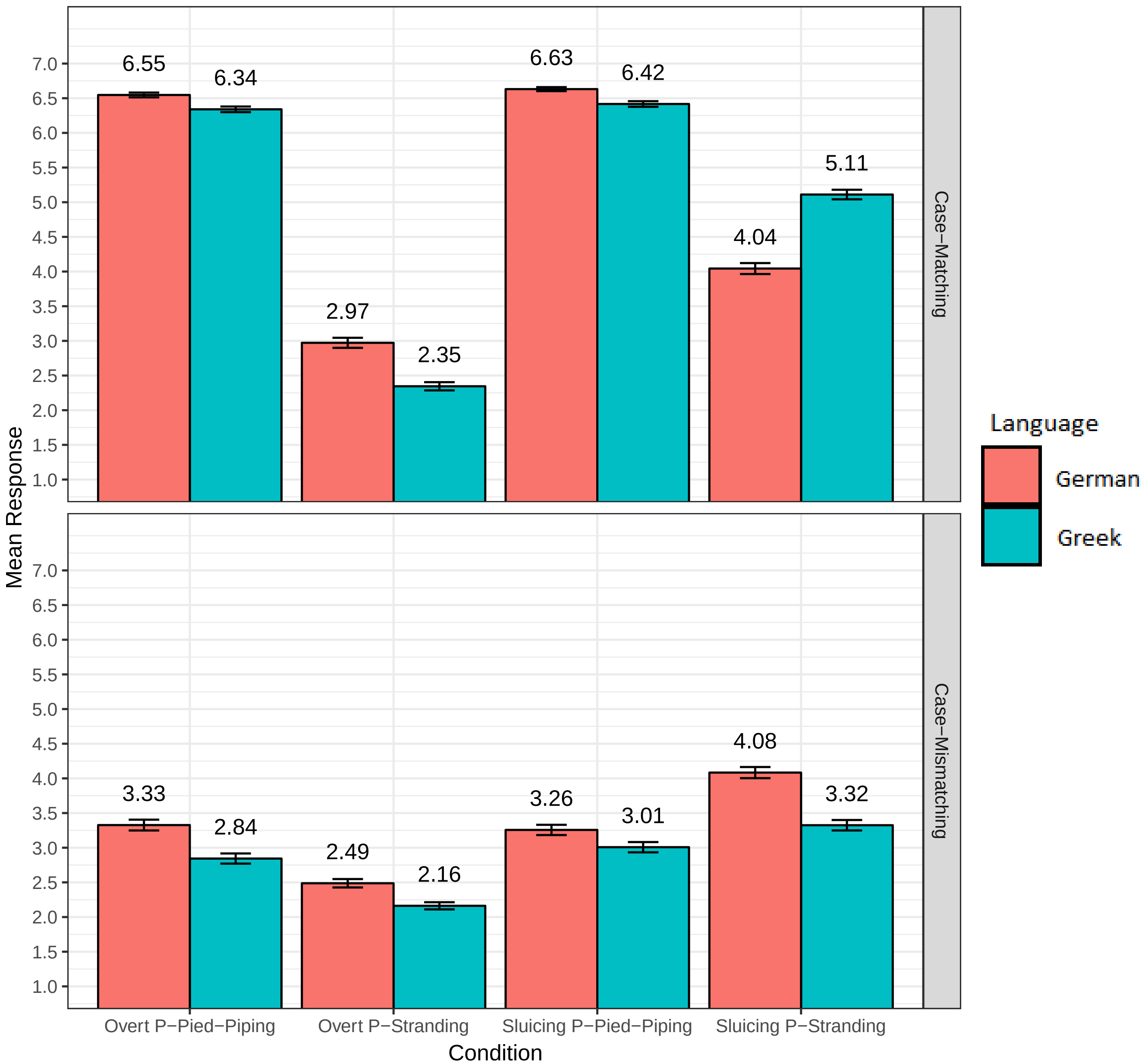


Table 2.27: Greek & German; Acceptability Ratings Numerical Summary

To save space, 'Non-Sluicing' conditions have been renamed 'Overt'.

	Condition	Language	Response	SD	SEM	95% CI
1	Overt Case Matching P-Pied-Piping	German	6.547	0.883	0.035	0.068
2	Overt Case Matching P-Stranding	German	2.972	1.845	0.073	0.143
3	Overt Case Mismatching P-Pied-Piping	German	3.327	1.980	0.078	0.153
4	Overt Case Mismatching P-Stranding	German	2.487	1.544	0.061	0.119
5	Sluicing Case Matching P-Pied_Piping	German	6.631	0.732	0.029	0.057
6	Sluicing Case Matching P-Stranding	German	4.043	2.014	0.079	0.155
7	Sluicing Case Mismatching P-Pied-Piping	German	3.257	1.879	0.074	0.145
8	Sluicing Case Mismatching P-Stranding	German	4.083	2.049	0.081	0.158
9	Overt Case Matching P-Pied-Piping	Greek	6.340	1.046	0.040	0.079
10	Overt Case Matching P-Stranding	Greek	2.345	1.528	0.059	0.116
11	Overt Case Mismatching P-Pied-Piping	Greek	2.844	1.902	0.073	0.144
12	Overt Case Mismatching P-Stranding	Greek	2.162	1.336	0.052	0.101
13	Sluicing Case Matching P-Pied_Piping	Greek	6.417	1.051	0.041	0.080
14	Sluicing Case Matching P-Stranding	Greek	5.111	1.783	0.069	0.135
15	Sluicing Case Mismatching P-Pied-Piping	Greek	3.008	1.909	0.074	0.145
16	Sluicing Case Mismatching P-Stranding	Greek	3.324	1.945	0.075	0.148

Table 2.28: Greek & German; Main Effects & Interactions

		Raw Scores		z-Scores		$Log10$-Scores	
	Factor	t-value	p-value	t-value	p-value	t-value	p-value
1	Case-Matching	69.41	< 0.0001	76.028	< 0.0001	59.923	< 0.0001
2	P-Stranding	52.19	< 0.0001	26.725	< 0.0001	12.335	< 0.0001
3	Sluicing	29.46	< 0.0001	15.880	< 0.0001	13.521	< 0.0001
4	Language	7.91	< 0.0001	13.767	< 0.0001	12.871	< 0.0001
5	Case-Matching*P-Stranding	46.68	< 0.0001	51.589	< 0.0001	38.338	< 0.0001
6	Case-Matching*Sluicing	5.13	< 0.0001	5.902	< 0.0001	4.454	< 0.0001
7	P-Stranding*Sluicing	27.23	< 0.0001	15.906	< 0.0001	11.716	< 0.0001
8	Case-Matching*Language	9.04	< 0.0001	12.347	< 0.0001	12.179	< 0.0001
9	P-Stranding*Language	1.75	> 0.0804	0.268	> 0.789	0.068	> 0.945
10	Sluicing*Language	6.98	< 0.0001	7.808	< 0.0001	6.765	< 0.0001
11	Case-Matching*P-Stranding*Sluicing	4.55	< 0.0001	4.743	< 0.0001	3.073	< 0.002
12	Case-Matching*P-Stranding*Language	5.73	< 0.0001	3.434	< 0.0006	0.564	> 0.572
13	Case-Matching*Sluicing*Language	7.82	< 0.0001	8.782	< 0.0001	7.499	< 0.0001
14	P-Stranding*Sluicing*Language	4.79	< 0.0001	5.612	< 0.0001	6.099	< 0.0001
15	Case-Matching*P-Stranding*Sluicing*Language	10.62	< 0.0001	11.037	< 0.0001	7.769	< 0.0001

fact that we did not find a main effect nor interaction of P-stranding with remnant and/or correlate complexity is also unfavourable for a pointer account which advocates that factors generally aiding in anaphora resolution should also significantly aid in sluicing resolution (Nykiel, 2013). In particular, we took this specific factor even further this time than in the original Greek study, comparing *which*-NPs with and without an overt NP to each other and to bare *wh*-remnants. By including equal numbers of both of these factors, we could check whether a particular combination of correlate-remnant information richness may be affecting the ease of P-stranding, as Nykiel (2013) appears to imply. We could thus check, for instance whether [complex correlate (*some buried treasure*) + simplex remnant (*which*)] or [complex correlate (*some buried treasure*) + complex remnant (*which buried treasure*)] had different sized effects on P-stranding amelioration, however, as evident in the results section, it appears that neither combination had any effect whatsoever, nor were they at all different from bare *wh*-remnants. As with all previous Greek studies, a pointer account thus leaves us with an incomplete account of the data.

This inadequacy of previous theories becomes even more apparent when we cast our eye on the case-mismatching results. Specifically, it would appear as though case-mismatching too is significantly better under sluicing compared to overtly, as much so, in fact, as P-stranding. Clearly this is not predicted by PF-Deletion, with or without the Fit Condition amendment, nor LF-Copying. It also comes as a surprise to a pointer account, whether such an account advocates for case-matching between correlate and remnant as an independently imposed sluicing pre-requisite or not. If it does, then this acceptability of case-mismatching under sluicing compared to non-sluicing conditions is in and of itself very perplexing; and if it does not, then it is unclear why case-mismatching is not always as acceptable as case-matching under sluicing, given there should be no structure being posited at the *e*-site. Furthermore, no other factors, such as remnant or connector type, appear to have any effect on this case-mismatching acceptability.

In conclusion, as we can see, none of the thus far proposed accounts of sluicing, or ellipsis more generally, can adequately explain these data.

2.6 General Discussion of both Greek and German data

Following these two sets of studies involving the overtly non-P-stranding languages of German and Greek, we can quite confidently conclude that P-stranding is not acceptable overtly but that it does appear to be significantly ameliorated under sluicing in both languages, whilst at the same time retaining its inferiority to P-pied-piping. Furthermore, thanks to the rich feature-marking of both languages, we were further able to investigate how case-mismatching compares to case-matching, both under sluicing and overtly, with results indicating that, again, sluicing does afford higher overall acceptability to this mismatch, albeit to different degrees in each language, with German case-mismatching under sluicing being rated more highly than its Greek equivalent (German: Mean: 3.67; SD: 1.96; Greek: Mean: 3.17; SD:1.93). Interestingly, P-stranding under sluicing was equally acceptable with and without case-matching in German, whereas there was a significant difference between the two in Greek. Finally, another interesting finding to point out is that, in both languages, overtly repeating unnecessary material does not appear to have a detrimental effect on acceptability, with Experiments 1 and 4 showing very close acceptability ratings for Case-matching P-pied-piping overtly and under sluicing, and Experiment 4, in particular, also finding no difference in results between *which*-NPs with an omitted NP vs. a pronounced NP. This is quite surprising for advocates of the 'repeated name penalty'.

We also concluded at the end of each experiment that none of the previously proposed general theories of sluicing, from PF-Deletion to a pointer account, appear to explain these data. Before we move on to propose a novel theory to fully explain both these and previous data, however, it is important to first consider in more detail how our results fare with respect to the previous explanations which have been proposed for other apparently PSG-defying languages (see e.g. section 1.3), either to try and keep results in line with a PF-Deletion account or not. In the sections below, we examine each of these ideas, as well as entertain some novel ones along the same lines, concluding that none can adequately capture the behaviour of either language, whether separately or in combination. Some of these have been briefly touched upon in previous sections, such as a copular or resumptive pronoun structure underlying P-stranding (1.3.2), or perhaps P-stranding being the result of a temporary grammatical illusion, whereas others are presented here for the first time.

2.6.1 P-stranding via an Alternative Source

Within a framework that posits structure at the *e*-site, the first popular explanation for the amelioration of P-stranding - or indeed any other violation - under sluicing is to simply propose an alternative structure is hidden or being inferred at the *e*-site; if this structure does not contain any violation, the issue is circumnavigated altogether (Abels, 2011; M. Barros et al., 2014; Rodrigues et al., 2009). The two most favoured alternative structures that have been proposed are, firstly, a cleft or copular structure, with or without a relative clause; and, secondly, resumption, as we have already seen. The exploration of the first option factored into the design of both Experiments 1 and 4 in the choice of case for the four case-mismatching conditions, whereas the second option is discussed here based on language-specific characteristics.

2.6.1.1 Sluicing equals Pseudosluicing

As a rule, in both Greek and German, cleft pivots, like copulas, must appear in Nominative case, as in (75) and (76).

(75) a. O Mitsos ine oreos.
Det.M.NOM *Mitsos*.M.NOM *is*.3SG *handsome*.M.NOM
'Mitsos is handsome.'

b. Ine o Mitsos pou ine oreos.
Is.3SG *Det*.M.NOM *Mitsos*.M.NOM *that* *is*.3SG *handsome*.M.NOM
'It is Mitsos that is handsome.'

(76) a. Der Hans ist schön.
Det.M.NOM *Hans* *is*.3SG *handsome*.M.NOM
'Hans is handsome.'

b. Es ist der Hans der schön ist.
It *is*.3SG *Det*.M.NOM *Hans* *who*.M.NOM *handsome*.M.NOM *is*.3SG
'It is Hans that is handsome.'

Thus, if one were to assume all cases of P-stranding under sluicing have in fact a hidden non-P-stranding copula at their core, one would expect to find this same default case-marking on the *wh*-remnant, whether we take the silent structure to be a bare cleft (78) or a cleft heading a relative clause (79). On the other hand, the object of a P must appear in Accusative in Greek, and in Accusative or Dative in German (depending on the P), with neither language's Ps licensing Nominative-marking. As a result, from a purely syntactic perspective, if what appear on the surface to be P-stranding sluices in fact contain a copular- and not a P-stranding source within the *e*-site, then we would expect to find that their bare sluice remnant can acceptably appear in Nominative, but not Accusative (or Dative); if, on the other hand, the *e*-site really does contain a P-stranding structure, then we would anticipate the opposite pattern.

(77) I Maria choreve me kapjon, ala de thimame
*Det.*F.NOM *Maria.*F.NOM *was.dancing.*3SG *with someone.*M.ACC *but* NEG *remember.*1SG
pjon /*pjos.
*who.*M.ACC *who.*M.NOM
'Maria was dancing with someone but I do not remember who.'

(78) I Maria choreve me kapjon, ala de thimame
*Det.*F.NOM *Maria.*F.NOM *was.dancing.*3SG *with someone.*M.ACC *but* NEG *remember.*1SG
*pjon /pjos itan.
*who.*M.ACC *who.*M.NOM *was.*3SG
'Maria was dancing with someone, but I do not remember who it was.'

(79) I Maria choreve me kapjon, ala de thimame
*Det.*F.NOM *Maria.*F.NOM *was.dancing.*3SG *with someone.*M.ACC *but* NEG *remember.*1SG
*pjon /pjos itan (aftos /ekinos) me ton
*who.*M.ACC *who.*M.NOM *was.*3SG *he.*M.NOM *that.one.*M.NOM *with Det.*M.ACC
opion choreve.
*which.*M.ACC *was.dancing.*3SG
'Maria was dancing with someone, but I do not remember who it was with whom she was dancing.'

To gather evidence in favour of either of these possibilities, our four case-mismatching conditions in both Greek and German always featured a Nominative-marked remnant; the Greek case-matching ones further featured an Accusative-marked remnant and the German ones a Dative-marked one. As seen in section 2.1.3, the results of the Greek study, in particular, clearly showed that these Nominative-marked remnants were significantly worse than their Accusative-marked counterparts, with case-mismatching always being found much less acceptable than case-matching across conditions (shown by a significant main effect of Case-matching). Although Nominative-marked remnants did appear slightly more acceptable under sluicing than did their overt moved counterparts, their acceptability was still nowhere near the Accusative-marked ones.

Even if we were to cast aside this case-matching issue in Greek, however, the results of our third experiment, involving contrast sluices, provides an additional argument against pseudosluicing being the source of Greek P-stranding. Specifically, as also mentioned in the outline of Experiment 3, clefting is not acceptable with contrast sluicing or 'else' modification (80, repeated below). As a result, we would expect that if P-stranded sluices could be reduced to pseudosluices, then these apparent 'P-stranding' structures should also be disallowed when contrast or 'else' modification sluicing is involved. As the results of Experiment 3 showed, however, this does not appear to be the case. Although there was an overall deterioration of sluicing acceptability in Experiment 3, this was across both P-pied-piping and P-stranding, with the latter still being significantly more acceptable under sluicing than overtly. As far as Greek is concerned, therefore, there are several clear arguments distinguishing P-stranding under sluicing from simply pseudosluicing.

(80) a. *I Maria chorepse me to GIANNI, alla de thimame pjos ALLOS itan.

*I	Maria	chorepse	me	to	GIANNI,	alla	de
Det.F.NOM	*Maria*.F.NOM	*danced*.3SG	*with*	*Det*.M.ACC	*Giannis*.M.ACC	*but*	NEG

thimame	pjos	ALLOS	itan.
remember.1SG	*who*.M.NOM	*else*.M.NOM	*was*.3SG

'*Maria danced with GIANNIS, but I do not remember who ELSE it was.'

b. *I Maria chorepse me pende NEAROUS, alla de thimame poses KOPELES itan.

*I	Maria	chorepse	me	pende	NEAROUS,	alla	de
Det.F.NOM	*Maria*.F.NOM	*danced*.3SG	*with*	*five*	*young.men*.M.ACC.PL.	*but*	NEG

thimame	poses	KOPELES	itan.
remember.1SG	*how.many*.F.ACC.PL.	*young.ladies*.F.ACC.PL.	*was*.3SG

'*Maria danced with five GUYS, but I do not remember how many GIRLS it was.'

(81) a. I Maria chorepse me to GIANNI, alla de thimame pjon ALLON.

I	Maria	chorepse	me	to	GIANNI,	alla	de
Det.F.NOM	*Maria*.F.NOM	*danced*.3SG	*with*	*Det*.M.ACC	*Giannis*.M.ACC	*but*	NEG

thimame	pjon	ALLON.
remember.1SG	*who*.M.ACC	*else*.M.ACC

'Maria danced with GIANNIS, but I do not remember who ELSE.'

b. I Maria chorepse me pende NEAROUS, alla de thimame poses KOPELES.

I	Maria	chorepse	me	pende	NEAROUS,	alla	de
Det.F.NOM	*Maria*.F.NOM	*danced*.3SG	*with*	*five*	*young.men*.M.ACC.PL.	*but*	NEG

thimame	poses	KOPELES.
remember.1SG	*how.many*.F.ACC.PL.	*young.ladies*.F.ACC.PL.

'Maria danced with five GUYS, but I do not remember how many GIRLS.'

With respect to German, this distinction is also present, although slightly less straightforward, at first glance. As we saw, in the German study both P-stranding with an Accusative-marked remnant and a Nominative-marked remnant scored equally highly under sluicing, with both being significantly better accepted than their overt counterparts. Furthermore, this amelioration of case-mismatching under sluicing was only present for P-stranding and not P-pied-piping, as indicated by an examination of the three-way interaction of Case-matching, Sluicing and P-stranding. As the results stand, therefore, it is currently impossible to rule out that German speakers may have been entertaining a cleft source for the Nominative-marked 'P-stranding' condition. The only way to check this would be to run a follow-up study in which we compare this condition to other, non-Nominative case-mismatching conditions. If, for instance, we find that Genitive-marked case-mismatching conditions are considered as acceptable as Nominative-marked case-mismatching conditions, then this would provide evidence against a copular source underlying the Nominative-marked conditions here. That having been said, however, the P-stranding remnants with Accusative-marking here are still incompatible with such a cleft source. As argued above, if pseudosluicing were the *only* source available to the parser for apparent P-stranding structures, then we would expect to find Accusative-marked bare remnants to be significantly less acceptable than their Nominative-marked counterparts, given the latter appear in the correct case for clefts. It is quite possible, therefore, that the P-stranded case-matching and case-mismatching sluices have different sources. This possibility will also make sense when we introduce a new way of thinking about sluicing in the coming chapter.

Of course, the next logical question would be why is it that German participants, in particular, appear able to entertain such a cleft source, whereas Greek participants do not? At present, I do not have a fully articulated answer to this question, however before entertaining explanations based on language-specific differences, I would first like to replicate these results not simply for these languages, but for others as well, using the same participant recruitment method as in German.

2.6.1.2 P-stranding by Resumption

In Wang (2007) and Boeckx (2008), the idea of a resumptive pronoun being behind the superficial amelioration of P-stranding under sluicing, similarly to that of island extraction, was put forth. Although there are several well-constructed counter-arguments to this logic in the theoretical literature (Merchant, 2001; Rottman & Yoshida, 2013), the experimental findings of AlShaalan & Abels (2019), with respect to Saudi Arabic, provide considerable evidence in its favour compared to these previous papers, showing practically that when a resumptive source is available, P-stranding is much more acceptable than when such a source is blocked. To recapitulate their experiments and results in brief, following our methods in Experiment 1, they too crossed the factors of P-stranding (±) with Sluicing (±). Given this is a non-case-marking language, Case-matching was dropped as a factor; instead, they compared some *wh*-phrases where an alternative *e-site* structure based on resumption was possible vs. others where resumption was not an option. They found that for those *wh*-phrases that allowed resumption within the *e*-site, P-stranding under sluicing was as acceptable as P-pied-piping, both overtly and covertly, whilst also being significantly more acceptable than P-stranding overtly. Importantly, for the other *wh*-phrases, those which block such resumption strategies, they found that P-stranding under sluicing was significantly less acceptable than both a) P-pied-piping in general and b) P-stranding when resumption is allowed. Even for these blocking *wh*-phrases, however, P-stranding was still significantly more acceptable under sluicing than overtly. That is to say, despite this experiment being the best evidence for an alternative source to P-stranding sometimes being available, this approach still cannot explain the full picture, even within the same language.

Given its popularity as a potential solution to P-stranding, let us briefly examine whether resumption could be the answer for our Greek and German results. For Greek in particular, although there exist forms of resumptive pronoun (mostly clitics), these have the property that they behave as regular pronouns and as such cannot actually constitute fillers for gaps (see also Chatsiou (2006) for more details, though cf. Alexiadou and Anagnostopoulou (2000)). As a result, in using these resumptive pronouns we cannot even salvage P-stranding overtly (82, 83), much less stipulate that they play a significant role covertly[25].

(82) * Pjon chorepse i Maria me afton /ton /ekinon ?
who.M.ACC *danced*.3SG *Det*.F.NOM *Maria*.F.NOM *with him*.M.ACC *him*.M.ACC *that.one*.M.ACC
'*Who did Maria dance with him?'

(83) I Maria chorepse me kapjon, ala den boro na po ...
Det.F.NOM *Maria*.F.NOM *danced*.3SG *with someone*.M.ACC *but* NEG *can*.1SG SUBJ *say*.1SG
'Maria danced with someone, but I couldn't say ...'

a. ... pjon
who.M.ACC
'... who.'

b. ... *pjon me afton /ton /ekinon chorepse.
who.M.ACC *with him*.M.ACC *him*.M.ACC *that.one*.M.ACC *danced*.3SG
'*... who she danced with him.'

[25]It has been argued that *wh*-XPs linked specifically to resumptive pronouns should appear in a default case - regardless of the pronoun's case (Merchant, 2001) - as they are no longer linked to a case-assigning base-position. In both the Greek and German examples we present, the judgements are the same regardless of *wh*-XP-case, hence to avoid an overly verbose set of glosses, the *wh*-XPs are shown only with the case matching form of resumptive pronoun.

c. ... *pjon chorepse me afton / ton /ekinon.
who.M.ACC *danced*.3SG *with him*.M.ACC *him*.M.ACC *that.one*.M.ACC
'*... who she danced with him.

The situation in German is similarly counter-productive for resumption as a resolution mechanism. Whether we take the regular pronoun *er* or the deictic *der*, neither of these appears to be allowed as an alternative filler for gaps, whether this be in direct (84) or embedded questions (85).

(84) * Wem hat die Maria mit ihm /dem getanzt?
who.M.ACC *danced*.3SG *Det*.F.NOM *Maria*.F.NOM *with him*.M.DAT *him*.M.DAT *danced*
'*Who did Maria dance with him?'

(85) Die Maria hat mit jemandem getanzt, aber ich weißnicht ...
Det.F.NOM *Maria*.F.NOM *has*.3SG *with someone*.M.DAT *danced but I know*.1SG NEG
'Maria danced with someone, but I don't know ... '

a. ... wem.
who.M.DAT
'... who.'

b. * ... wem sie mit ihm /dem getanzt hat.
who.M.DAT *she*.F.NOM *with him*.M.ACC *him*.M.ACC *danced has*.3SG
'*.. who she danced with him.'

As such, it would appear as if resumption is not an acceptable option for P-stranding amelioration under sluicing, either in Greek or German. That having been said, we must concede one point. If P-stranding under sluicing can indeed be explained via an alternative source with no further changes being necessary, then both this overt alternative source and P-stranding under sluicing should show very similar acceptability ratings. This, for instance, was apparently the case with Saudi Arabic. Following this line of thinking, if the reason we find P-stranding under sluicing to be less-than-perfect in both Greek and German (compared e.g. to P-pied-piping under sluicing) is that their overt alternative source is similarly less-than-perfect, then this would serve as a convincing argument in favour of such an alternative source. Given we do not have any experimental acceptability ratings for such alternative sources beyond our informants, we must logically allow for this possibility. If, however, we find that overtly they are completely unacceptable, as reported in the above examples, then it would be nonsensical to reduce P-stranding under sluicing to such sources. Furthermore, if P-stranding under sluicing can be explained solely by such alternative sources, then we are still left with the question of why Saudi Arabic still finds these sluicing conditions acceptable, albeit to a smaller degree, even when these alternatives are overtly much less acceptable in comparison.

2.6.2 P-Stranding as a Grammatical Illusion

Another suggestion that could be made is that this apparent acceptability of P-stranding under sluicing is nothing but a grammatical illusion, similar, for instance, to that of agreement attraction (Bock & Miller, 1991; Eberhard, Cutting, & Bock, 2005; Pearlmutter, Garnsey, & Bock, 1999). The idea behind a grammatical illusion is that, at least temporarily, what is essentially an ungrammatical utterance may be perceived as grammatical, for various reasons. In agreement attraction, for example, a verb may agree with an inappropriate NP instead of its subject on a marked feature, such as number (e.g. (86), from Bock and Miller (1991), their (1)). This has been interpreted as the intervening NP's number-marking somehow licensing agreement on the verb. This effect occurs both with proximally intervening and disagreeing NPs (Francis, 1986; Quirk, Greenbaum, Leech, & Svartvik, 1985), but also with non-local nouns, for instance when the verb of a relative clause agrees with the clause head instead of its own subject (Bock & Miller, 1991; Franck, Lassi, Frauenfelder, & Rizzi, 2006; Franck, Soare, Frauenfelder, & Rizzi, 2010), and has been found in both production and comprehension (Eberhard et al., 2005; Häussler & Bader, 2009).

(86) The time for fun and games are over.

This type of illusion is not restricted solely to number features, however, but has also been shown to locally affect case-marking (Bader, Meng, & Bayer, 2000; Sloggett, 2013), which is of more interest to us here.

Following this train of thought, one could argue that we may be dealing with something similar here, too. When participants see the bare remnant directly following the verb which introduces the sluice, e.g. the verb *know* in *Maria danced with someone, but I don't know who*, what is intended to be a P-stranded remnant may be being more leniently interpreted as the direct Accusative-marked object of this verb, thereby giving the reader a fleeting sense of grammaticality.

To test for this logical possibility, all Greek experiments presented here were carefully constructed so that half of the stimuli contained embedding verbs which patterned with an Accusative complement (as at least one of its c-selection options), whereas the other half could not pattern with Accusative, instead c-selecting for a Genitive-marked NP or a full CP. The analysis of the results found no significant difference between the two groups of results in any experiment, with the factor of Embedding Verb having no main effect or significant interactions with any of the fixed effects factors in any experiment. As such, one can quite confidently conclude that the ratings of P-stranding under sluicing are not simply due to some short-lived semblance of acceptability based on the embedding verb's syntax.

For the German study, unfortunately due to an error during stimuli creation, these were not equally split between embedding verbs with [+Dat.] and [-Dat.] c-selection options, therefore this analysis was not possible.

2.6.3 Is Stranding the P the same as Overtly Dropping it?

A further theory not mentioned previously, but which could be argued for, is that P-stranding is in fact an instance of *P-dropping*, a phenomenon where certain locative NPs normally appearing within a PP are allowed to appear without that P, effectively *dropping* it. One of the more successful theories proposed to explain this is that of the bare NP representing a full PP reading, i.e. a type of phrasal lexicalisation having occurred (Caha, 2011; Collins, 2007; Ioannidou & Den Dikken, 2009; McCawley, 1988). This approach would also make sense from the point of view of the Fit Condition, i.e. if the full PP can be replaced by the bare NP whilst maintaining the same meaning and syntactic function, then it would indeed be able to 'fit' in the full PP's place in the antecedent, making it an acceptable remnant. It should be noted that Modern Greek too, like a number of other languages, exhibits this same phenomenon of overt P-dropping (Caha, 2011; Ioannidou & Den Dikken, 2009), thus making it tempting at first glance to consider P-stranding simply an extension of this phenomenon, at least for Greek. However, there are a few crucial points to consider which clearly differentiate the two.

First of all, as also mentioned in Ioannidou and Den Dikken (2009), acceptable P-dropping in Greek is lexically restricted to certain nouns very commonly used for stereotypical locations, such as for instance *beach, seaside, home, bar, nightclub, bouzoukia*[26], amongst others. This form of P-dropping is not acceptable with any other lexical items, such as indefinite quantifiers, *wh*-elements, proper names or a number of other elements which may, however, appear as sluice remnants (88). Secondly, there are restrictions on the types of PP for which this is an available option. Specifically, dropping the preposition is solely applicable for locative PPs introduced by the preposition 'se' (*to*), whereas P-stranding also occurs under sluicing with a number of other prepositions in the correlate PP (e.g. 'jia' (*for*), 'me' (*with*), 'apo' (*from)*, 'kata' (*against*) a.o.) (89), with no significant difference found between the four we examined in any of our experiments. Finally, instances of P-dropped NPs must appear without adjectives or determiners, as opposed to P-stranded remnants. Examples ((87) through (90) compare P-dropping to P-stranding instances, allowing us to conclude

[26]A traditional string instrument, the noun used metonymically to refer to a type of night club where the instrument is typically played live on stage.

that P-stranding cannot simply be another case of P-dropping in this form, at least as far as Greek is concerned.

(87) a. Ta pedja tha pane (s- tin) paralia.
*Det.N.NOM.PL. children.N.NOM.PL. FUT. go-3PL. to- Det.*F.ACC *beach.*F.ACC
'The children will go to the beach.'

b. Ta pedjia tha pane se kapja paralia, ala de thimomaste pja.
*Det.N.NOM.PL. children.N.NOM.PL. FUT. go-3PL. to some.*F.ACC *beach.*F.ACC *but* NEG *remember-1PL. which.*F.ACC
'The childrean will go to a certain beach, but we do not remember which (one).'

(88) a. Ta pedja tha pane *(s- ti) jiajia.
*Det.N.NOM.PL. children.N.NOM.PL. FUT. go-3PL. to- Det.*F.ACC *grandma.*F.ACC
'The children will go to Grandma.'

b. Ta pedjia tha pane se kapja jiajia (tous), ala de thimame pja.
*Det.N.NOM.PL. children.N.NOM.PL. FUT. go-3PL. to some.*F.ACC *grandma.*F.ACC *(theirs) but* NEG *remember.*1SG *which.*F.ACC
'The children will go to some Grandma/one of their Grandmas, but I do not remember which (one).'

(89) a. Ta pedjia tha erthoun *(apo *(tin)) paralia.
*Det.N.NOM.PL. children.N.NOM.PL. FUT. come-3PL. from Det.*F.ACC *beach.*F.ACC
'The children will come from the beach.'

b. Ta pedjia tha erthoun apo kapja paralia, ala de thimame pja.
*Det.N.NOM.PL. children.N.NOM.PL. FUT. come-3PL. from some.*F.ACC *beach.*F.ACC *but* NEG *remember.*1SG *which.*F.ACC
'The children will come from some beach, but I do not remember which (one).'

(90) a. Ta pedjia tha pane *(s- *(-ti)) megali
*Det.N.NOM.PL. children.N.NOM.PL. FUT. go-3PL. to- Det.*F.ACC *big.*F.ACC

paralia.
beach.F.ACC
'The children will go to the big beach.'

b. Ta pedjia tha pane se kapja megali
Det.N.NOM.PL. children.N.NOM.PL. FUT. go-PL. to some.F.ACC *big*.F.ACC
paralia, ala de thimame pja megali paralia.
beach.F.ACC *but* NEG *remember*.1SG *which*.F.ACC *big*.F.ACC *beach*.F.ACC
'The children will go to some big beach, but I do not remember which big beach.'

With respect to German, this form of P-dropping appears to be allowed in Kiezdeutsch, a particular multi-ethnic variety of German spoken mainly by young people in urban areas. Outside of this variety, P-dropping is severely limited. Specifically, it appears only informally and only with public transport stop NPs .

(91) Da fahren Sie bis (zur) Friedrichstraße
PRT *go* *you*.2RP *until to.Det*.DAT *Friedrich.street*
'You have to go to Friedrichstraße, then.'

Even in Kiezdeutsch, however, this phenomenon is limited in certain respects. Firstly, the P complement is almost always a bare NP, i.e. the Determiner must also be dropped; secondly, outside of specific directional verbs, such as *fahren* ('ride/go') and *gehen* ('go'), and the verb *to be*, it seems to interact with the verbs it appears with, semantically reducing or 'bleaching' them, with these verbs losing their full lexical content Wiese (2009, p.22). A more detailed explanation of this phenomenon is outside the scope of this discussion, with the reader being referred to Wiese (2009) for a more detailed analysis on the German variety of Kiezdeutsch and its use of a different verb pattern to standard German. Suffice it to say, however, that even if all our participants spoke this variety of German, in itself highly unlikely, P-dropping would still not be compatible with our experimental items given their use of standard German verb forms and meanings. As such, we can quite confidently conclude that what we are seeing in the German study is not simply an instance of P-dropping as exhibited in this 'youth language'.

2.6.4 Greek and German are *not* 'PSG-exceptions'

Although there may be other explanations for P-stranding under sluicing in overtly non-P-stranding languages, these are the most well-known, to the extent of our knowledge. As is evident from the above sections, it does not appear as though any of these can clearly and neatly explain both the Greek and German data we have found. Furthermore, given none of the existing theories of sluicing can explain them fully either, this leaves us with an intriguing puzzle to solve. In the following chapter, I will provide a solution to this riddle, proposing a theory which combines PF-Deletion, in terms of structure posited at the *e*-site, together with a sentence processing approach that shall treat sluicing as a form of *noisy channel*. As shall be explained more thoroughly in due course, the parser shall be proposed to act as a probabilistic inference machine in the absence of overt evidence, calculating and inferring the best possible structure as a continuation for the remnant based on prior experience and available evidence to hand. This processing mechanism and the costs it entails shall be proposed to have a direct effect on the structure's acceptability, as many others have also previously argued (Hofmeister, 2007; Hofmeister, Casasanto, & Sag, 2014; Hofmeister et al., 2013; Sprouse, 2008; Sprouse, Wagers, & Phillips, 2012), a.m.o.). In this way, I shall attempt to not only explain the data presented thus far, but also generate easily testable predictions for all other languages. Simply put, this theory will present a very distinct and important advantage over any other existing one as a) it shall attempt to provide a unified theory of real-world sluicing behaviour for all languages, whether these have been previously proposed to comply with or defy the PSG; and b) as a result, it will lay out simply verifiable experimental predictions for each language for future studies to provide evidence for or disprove, as any scientific hypothesis should. Given the lack of large-scale and controlled experiments in sluicing in almost any other language, this provides further incentive for future such studies to confirm or disprove these predictions.

Chapter 3

A Noisy Channel Model of Sluicing

3.1 Introduction

In the first chapter, we had a look at the various proposed theoretical underpinnings of sluicing more generally and P-stranding under sluicing more specifically. A brief overview was given of the cross-linguistic scene with respect to the acceptability of P-stranding under sluicing in overtly non-P-stranding languages. We saw that a variety of anecdotal evidence exists pointing towards this phenomenon being in fact more wide-spread than one might originally imagine or predict, however precious little adequate experimental evidence exists to support it and help identify a cohesive cross-linguistic explanation. Indeed, despite these anecdotal data being presented several times in the literature, no theory has thus far provided a successful unified approach to them.

In the second chapter, four highly controlled, large-scale acceptability judgement studies were presented providing the first experimental evidence of its kind for sluicing in Greek and German. The results showed that although P-stranding was evidently unacceptable in overt interrogative clauses, it became significantly more acceptable under sluicing in both languages, whilst at the same time remaining less acceptable than P-pied-piping both overtly and under sluicing. Evidence was additionally provided for case-mismatching also being more acceptable under sluicing than it is overtly. Through careful experimental design and the examination of each existing theory's predictions, we were able to show that no existing approach, whether theoretically or processing oriented, could account for this dataset, leaving us with a puzzling mystery.

In this chapter, I will present the background and more detailed concept of a sentence processing model which will then form the basis of my explanation for these intriguing data. The model in question is that of linguistic Bayesian estimation through the medium of a noisy channel, based, more generally, on the concepts of anticipation or prediction, surprisal and probabilistic estimation and, more specifically, on concepts from J. Hale (2001); Levy (2008b); Shannon (1948) and Gibson, Bergen, and Piantadosi (2013). This model will be used alongside a theoretical analysis of sluicing as a potential explanation for the experimental data presented in chapter 2, as well as to generate predictions for the behaviour of other languages with respect to sluicing. It will also be used to shed more light on the experimental data we will see in the next chapter.

3.2 Anticipation, Bayesian Estimation & Noise

3.2.1 The What; Anticipation and/or Predictability

Over a number of years, time and again it has been shown that, when interpreting linguistic material, as comprehenders we do not wait until the end of a sentence in order to begin interpreting it. Instead, we begin to rapidly segment utterances and identify meanings and relationships in the structure immediately upon hearing the first word. For instance, whilst reading a sentence containing an anaphor, there is evidence that we immediately take into consideration syntactic restrictions, such as binding conditions, to help identify the anaphor's meaning before reaching the end of the sentence (Kazanina, Lau, Lieberman, Yoshida, and Phillips (2007); Runner, Sussman, and Tanenhaus (2003); Sturt (2003), a.o., though cf. Badecker and Straub (2002)). There are various competing theories which have been proposed over the decades as to what this parsing process looks like exactly, i.e. whether syntactic structure is the first building block to be immediately erected (Frazier & Fodor, 1978), whether it is the properties of different words that govern this parsing process (Ford, Bresnan, & Kaplan, n.d.), whether semantics bears the lion's share (Pritchett, 1988; Tanenhaus et al., 1989), among many other approaches. Regardless of the approach taken, however, the efficiency of this process is such that in order for comprehenders to maximise the potential for interpretation in the minimum amount of time, it has been repeatedly argued that we must anticipate upcoming material on a plethora of linguistic levels - lexical, syntactic, semantic, phonetic - based on already seen material (G. Altmann and Steedman (1988); Arnon and Snider (2010); Demberg, Keller, and Koller (2013); Kleinschmidt and Jaeger (2015); W. Marslen-Wilson (1973); W. Marslen-Wilson and Tyler (1975); W. D. Marslen-Wilson (1975); W. D. Marslen-Wilson and Welsh (1978), among many others, though cf. Van Petten and Luka (2012) for discussion). This concept of communication efficiency can trace its origins back to Shannon (1948), whose seminal ideas on communication arguably founded the entire field of information theory, i.e. the study of how we quantify, store, convey and interpret information in all its forms. We will return back to some central concepts from information theory in due course. For now, it is this anticipation which shall be the focus of this section.

Evidence for this anticipation in language has been presented in the literature in a variety of experimental contexts; in eye-tracking studies, for instance, it takes on the form of anticipatory eye movements, with participants either skipping or focusing less on predictable words in reading tasks (Balota, Pollatsek, & Rayner, 1985; Demberg & Keller, 2008; Demberg et al., 2013; Frank & Bod, 2011; Rayner, 1998; Rayner, Binder, Ashby, & Pollatsek, 2001; Rayner & Well, 1996; Smith & Levy, 2013; Staub, 2015), or reflexively looking in visual world experiments only towards images which are semantically congruous with preceding context (G. T. Altmann & Kamide, 1999) (see Staub (2015) for overview). For example, in a situation where participants see 4 images on screen, those of a little girl, a man, a tricyle and a motorbike, when hearing the sentence 'The little girl will ride the...', they have been shown to look more often and for longer periods of time towards the image of the tricycle in advance of actually hearing the word in the sentence. On the other hand, when the participant hears 'The man will ride the...', he or she appears to look in anticipation towards the image of the motorbike, instead. In studies measuring reaction times, for naming (Balota et al., 1985; Monsell, Doyle, & Haggard, 1989), lexical decision (Arnon & Snider, 2010; Forster & Chambers, 1973; Murray & Forster, 2004)) and semantic classification tasks (Forster & Hector, 2002; Forster & Shen, 1996), participants have also been shown to be significantly faster for more predictable vs. less predictable words, based on context. This reaction time boost does not appear solely for words which are predictable, but also extends to larger phrases or parts of speech (POS) (Arnon & Snider, 2010). Further evidence comes from electroencephalography (EEG) experiments, where there is at least one event-related potential (ERP), the N400[1], which has been repeatedly associated specifically with semantically unexpected material, across written, spoken and signed words and pseudowords (Frank, Otten, Galli, & Vigliocco, 2013; Kutas & Hillyard, 1980, 1984) (see Kutas and Federmeier (2011) for review). Although in more recent years the exact nature of this response (and of other related ERPs) has become less clear than originally proposed, with some contention as to whether its amplitude is affected primarily by spreading semantic and/or lexical association vs prediction, there is at least some evidence pointing towards the N400 being influenced by anticipation when isolated from association, showing smaller amplitudes when a word is contextually predictable vs. when it is not, when lexical association is held constant (E. F. Lau, Holcomb, & Kuperberg, 2013). Moving towards the interface with production, in conversational contexts, it has also been argued that the reason we do not (usually) talk over each other is because we can predict when our turn will come based on the discussion's lexical and syntactic content (De Ruiter, Mitterer, & Enfield, 2006; Garrod & Pickering, 2015; Magyari & de Ruiter, 2012; Sacks, Schegloff, & Jefferson, 1978; Stivers et al., 2009).

[1] A negative deflection localised over the centro-parietal area, peaking at approximately $400ms$ post-stimulus onset.

Anticipation of upcoming material can also be witnessed in situations where it is more of a hindrance than a help. Consider, for instance, the famous garden-path effect during sentence comprehension (92). In temporarily ambiguous phrases such as this one, anticipation of the more frequently encountered phrase structure may - initially at least - (partially) block any other interpretation of the phrase. For instance, in English the word *raced* appears in general more frequently as the past tense of the verb *race*, as opposed to its past participle. As a result, comprehenders are more likely to interpret the structure in (92) as a subject followed by a past-tense verb, rather than a past participle. This is also the simpler of the two parses. In this case, when they reach the real verb, *fell*, they are forced to reanalyse or reinterpret the sentence in order to incorporate this unexpected word[2].

(92) The horse raced past the barn fell.

This reanalysis has been argued to lead to increased processing difficulty, expressed in both slower reading times at (or just after) the disambiguating verb, '*fell*', (Ferreira & Clifton Jr, 1986; Garnsey et al., 1997; MacDonald, Just, & Carpenter, 1992; Spivey-Knowlton, Trueswell, & Tanenhaus, 1993) as well as lower overall comprehension accuracy (Ferreira, Christianson, & Hollingworth, 2001; Ferreira & Patson, 2007).

Rather intriguingly, this process of anticipation or expectation appears, furthermore, to be largely flexible and interactive in its effects. It has been argued that different grains of linguistic representation (from morpheme to larger phrase) and even levels (semantic, syntactic, lexical, phonetic etc.) affect one another. For instance, in example (93), the most anticipated continuation would be the word *kite* (Federmeier, Wlotko, De Ochoa-Dewald, & Kutas, 2007). In an ERP study utilising such constraining contexts, DeLong, Urbach, and Kutas (2005) found that when the structure in (93) was presented in written form to participants either with the determiner *a* at the end vs. *an*, there was a larger negativity associated with *an* compared to *a*. Since *an* can only occur with words beginning with a vowel, this makes it inconsistent with the predicted word *kite*. This was considered as indication of preceding semantic context not only encouraging anticipation of the semantics of a word, but also its phonological and orthographic content. The relationship of anticipation and input appears to be even more complex when we consider evidence that it is not just the structures, words and contexts that we encounter, i.e. the bottom-up input, which affect our predictions, but also our previous expectations of a situation, i.e. our top-down knowledge, which can affect how we perceive bottom-up input (Kuperberg & Jaeger, 2016).

[2]How exactly this is done, i.e. whether the parser serially entertains only one sentence interpretation at a time and is then forced to discard it and reanalyse the whole sentence anew (Ferreira & Clifton Jr, 1986; Frazier, 1978), or whether it holds in parallel multiple possible parses which are simply weighted based on frequency and then re-arranged when bottom-up input forces the parser to discount the highest weighted one, this is a question for another time. By using one term or the other, I do not intend to reflect any commitment towards one theory or the other, although this is something which will be touched upon again later.

This is something which we will return to in more detail later on in this chapter, as well as when examining the data in chapter 4.

(93) The day was breezy so the boy went outside to fly a ...

An issue which should be mentioned at this point is that depending on the authors, the concept of anticipation or expectation can sometimes be considered the same as that of prediction, but other times deliberately distinguished from it (e.g. Van Petten and Luka (2012)). There is, furthermore, a conceptual distinction between predictive pre-activation and predictive commitment (E. F. Lau et al., 2013). For the intents and purposes of this chapter, it is not necessary to distinguish between these concepts, we will instead simply be discussing prediction or anticipation as interchangeable and their theoretical antithesis, surprisal.

3.2.2 The How; Frequency, Probability & Surprisal

3.2.2.1 Defining Predictability through Surprisal

Having established that anticipation or predictability of upcoming material is widely regarded to have a measurable effect on the way with which we parse linguistic input, we inevitably come to the question of how anticipation or predictability in and of itself should practically be defined or measured. From the original account of efficient communication proposed by Shannon (1948) to the more statistically oriented J. Hale (2001); Levy (2008a, 2008b) and Gibson et al. (2013), the key uniting and recurring factor in defining predictability is probability of occurrence. Probability in this sense first appears in two separate definitions of the concept of *surprisal*, firstly in Shannon (1948) and then again slightly differently in J. Hale (2001), with the two being unified later on by Levy (2008a). Surprisal, as the name might suggest, could be considered more informally as the opposite of predictability, an indication of how *unpredictable* an upcoming word or other linguistic unit is, given the material seen so far.

Beginning chronologically with Shannon (1948), the concept of predictability is first formally defined in reverse through the idea of surprisal. The way this works is linked to how Shannon proposed that information is conveyed in any communication setting. Each word in a string or sentence is considered to bring a certain amount of new information into play, tied to how expected or not the word is within the context it appears in. The more expected a word is, the less information it carries, and vice versa, the more new information the word carries, the more unexpected it is, and therefore, by extension, the more *surprising*. The more surprising a word is, the more cognitive effort it requires to be processed. In mathematical terms, given a word w_i appearing after a string of words

$w_{1...i-1}$, its surprisal can be quantified using the equation in (3.1) as the negative logarithm of this word's probability, conditional upon the string encountered thus far.

$$surprisal(w_i) = -logP(w_i|w_{1...i-1}) \tag{3.1}$$

In the interest of brevity, I will continue to use the term 'word' when discussing surprisal, however the reader should infer that, unless expressly indicated, 'word' is simply a placeholder that can be replaced by any other linguistic unit of input desirable; the equation in (3.1) can, thus, be used to represent the surprisal of an upcoming POS, given the POS context thus far, i.e. to predict purely syntactic effects; to represent an upcoming phoneme, given the phonemic/phonetic context thus far, i.e. to predict purely phonetic effects; and so on and so forth (though cf. Frank and Bod (2011)).

In his 2001 paper, Hale also discusses the surprisal of an upcoming word as the cognitive effort required to process it, albeit restricted to the level of syntactic processing. The surprisal of a given point i in a sentence is defined as 'the total probability of structural options that have been disconfirmed at that point' (p.1). This value is proposed to be calculable using a probabilistic Earley parser based on Stolcke (1995), a parsing mechanism originally designed to aid in speech recognition systems. When applied to the level of syntax, this parsing algorithm works on a word-by-word basis, at each word generating parallel top-down predictions or hypotheses about the upcoming material based on the context seen thus far, with bottom-up material confirming or disconfirming these hypotheses. Two main principles this model assumes and which are of interest here are that firstly, sentence processing is considered to be 'eager'. That is to say that Earley parsers work in a top-down fashion, rapidly generating predictions about upcoming material prior to bottom-up input, based on previous knowledge. Secondly, the pure frequency with which structures occur affects their probability of occurring, with greater frequency affording larger probability. If the probability assigned to a number of alternative structures is much greater than that of the actual structure at hand, then disconfirming these alternatives (through bottom-up input), which together comprise a large amount of probability, leads to a parsing slow-down. This is, for instance, the explanation this model proposes for garden-path effects, as opposed to an explanation based on e.g. structural reanalysis.

Although the above are somewhat simplified explanations of these two surprisal accounts, the main point to keep in mind is that probability of upcoming material is tied to previous knowledge, in terms of a) frequency of occurence (J. Hale, 2001) and b) its sentential context (Shannon, 1948).

These two slightly different views of surprisal are unified in 2008 by Levy. Expanding upon Shannon's information theoretic account of surprisal, Levy (2008a) unites it with Hale's definition of surprisal, showing that mathematically, the two are essentially equal and can generate accurate predictions regarding the cognitive effort required to process various words. The way this is done is by focusing on a related probabilistic concept of word predictability put forth by Shannon, namely *entropy*. According to Levy, we can assume that parsing any given sentence S initially involves placing a probability distribution over T, the set of all possible structures in that language, in order to eventually figure out the specific structure that S corresponds to. Starting out, we begin with a large degree of uncertainty (H) regarding the upcoming sentence (Shannon, 1948). However, moving forward, the parser incrementally updates this probability distribution over T to more accurately reflect new knowledge or word information (in Shannon's sense) obtained from each new word or sentence segment[3]. By updating this distribution continuously, the uncertainty about upcoming material also slowly shrinks. The distribution placed for an upcoming word w_i is based upon all previous words encountered from w_1 up until w_{i-1}, as also posited by Shannon. To this contextual influence upon the distribution, Levy furthermore adds the wider extra-sentential context (world-knowledge, discourse context etc.). The processing difficulty incurred by a word w_i is a direct result of updating this probability distribution from before to after seeing w_i and of shrinking uncertainty regarding the full sentence's distribution. The difference between the old distribution up until w_{i-1} and the new distribution updated at w_i, or the degree of reduction of uncertainty (ΔH), is known as *relative entropy* (Levy, 2008a) or *entropy reduction* (J. Hale, 2001). In statistical terms, this relative entropy is also known as the *Kullback-Leibler divergence*. The larger this relative entropy, i.e. the larger the difference between the old and new distributions, the larger the processing difficulty associated with parsing w_i. In other words, the less likely a new word is to occur, based on the previous words plus the wider context, the harder it is to integrate. This concept is shown mathematically to be exactly equivalent to Hale's surprisal. In this way, Levy unites the two theories giving a more general surprisal account of word processing, based on the narrower and wider context. 'Surprisal is thus minimized (goes to zero) when a word *must* appear in a given context (i.e., when $P(w_i|w_{1...i-1}, CONTEXT) = 1$), and approaches infinity as a word becomes less and less likely' (emphasis not in original)(p.5). One final, very important contribution by Levy (2008a) is the concept that the difficulty in processing an upcoming word is directly and *exclusively* linked to that word's surprisal. That is to say, surprisal is not considered simply an additional measure of calculating the cognitive effort required to parse a word or structure above and beyond the word or structure's inherent complexity. Instead, it is considered the *only* measure of cognitive effort. Levy thus considers cognitive effort to be exclusively linked to surprisal, i.e. the word's conditional probability, regardless of the linguistic unit's complexity. Surprisal in this way functions as a 'causal bottleneck'

[3]As we do above, for simplicity's sake, Levy assumes that these updates happen after each word. That having been said, one could argue that they occur at larger or even smaller intervals, perhaps even within the same word. Updates occurring within the same word, after each morpheme, could for instance predict a slow-down in processing when multiple words are compatible with the intra-word string thus far, such as when we observe competition interference effects whilst processing words that have a large neighbourhood density (Vitevitch & Luce, 1999).

between representations and observable input. Another advantage of the surprisal model is that it does not appear to show a processing granularity bias, or in other words, the model does not appear to make different predictions based on whether we are focusing on the morpheme vs. the word vs. the larger syntactic or semantic structure. Given the model is based entirely on what is most likely to follow the observed structures, the only way that the model's predictions would change based on granularity of observed and upcoming elements would be if there were indeed different probabilities associated with these different levels of granularity. That is to say, if there is no difference between the event likelihood associated with e.g. a word vs. the event likelihood associated with a larger syntactic structure that the word is simply a part of, then there will also be no difference associated with granularity reflected in the model.

The above theoretical considerations of probability and uncertainty in the parser could also be expressed in terms of Bayesian conditional probabilities, as both Levy (2008a) and Kuperberg and Jaeger (2016) point out. Bayes's rule is given in (3.2) and can be read as: the posterior probability of event A happening given event B ($P(A|B)$) is equal to the probability of event B happening given event A ($P(A|B)$) multiplied by the prior probability of event A happening anyway ($P(A)$) divided by the prior probability of event B happening anyway ($P(B)$).

$$P(A|B) = \frac{P(B|A) * P(A)}{P(B)} \tag{3.2}$$

What this means is that whenever we are calculating the probability of something, we always go into this calculation having a certain amount of prior knowledge about the events at hand, which helps to guide us. This knowledge allows us to make prior hypotheses about these events occuring, known as prior *beliefs*. This is similar to the way the parser might entertain hypotheses about upcoming material based on previous exposure (in terms of frequency) to various words, syntactic structures or semantic schemata in similar contexts. Upon receiving new input, this prior belief is then confirmed or disconfirmed and is translated into the posterior belief. This new posterior belief then acts as the prior belief for the next set of calculations and so on and so forth. Based on the definitions above, we could assume that this cycle of calculations occurs at each word and is used to determine the next word in a string, based on the previous words in that string. This shift is known as *belief updating*, and the degree to which our prior belief changes to become the posterior belief is known as *Bayesian surprise* (Doya, Ishii, Pouget, & Rao, 2007), or in the terms defined earlier on, this would be considered the relative entropy or entropy reduction.

Finally, it should be pointed out that this mathematical equivalence with Bayesian estimation also has an important effect on whether the surprisal account is to be considered a serial- or parallel-processing account. Under a serial-processing approach, the parser is considered to generate only one prediction at a time, starting with the most probable one (see below on how this is done practically), and holding it until this prediction is either confirmed or disconfirmed through bottom-up input. If it is disconfirmed, the parser then reanalyses and turns to the next most probable continuation and so on and so forth (Van Gompel, Pickering, Pearson, & Liversedge, 2005; Van Gompel, Pickering, & Traxler, 2001; Van Petten & Luka, 2012). Such approaches include, for instance, the Tuning hypothesis, which predicts that syntactic ambiguities are resolved by the parser always choosing the most frequently occurring structural option (Mitchell, 1994; Mitchell, Cuetos, Corley, & Brysbaert, 1995). Under a parallel-processing approach, on the other hand, the parser is considered to compute an array of probable continuations in parallel, each with a different weighting relative to their probability (De Ruiter et al., 2006; Staub, Grant, Astheimer, & Cohen, 2015; Wlotko & Federmeier, 2012). This does not necessarily mean that all possible continuations are predicted exhaustively in parallel, but rather that there are simply multiple probable continuations that are considered until bottom-up input disconfirms all but one (though cf. Crocker and Brants (2000); Jurafsky (1996); see Traxler (2014) for discussion). When one is disconfirmed, it is assumed that the remaining parallel continuations are re-ranked or re-weighted, with this being argued to be the main source of cognitive effort (Crocker & Brants, 2000; P. Gorrell, 1989; P. G. Gorrell, 1987; Jurafsky, 1996; Narayanan & Jurafsky, 2002).

Given the surprisal account's equivalence to Bayesian estimation, on the one hand, and how Bayesian estimation, on the other hand, considers multiple predictions in parallel, each with a different weighting, it has been argued that the general surprisal account should also be considered a parallel-processing account. However, it should be noted that the surprisal approach does have one big difference to both serial-parsing and other parallel-parsing accounts with respect to how it predicts ambiguity to affect the parser. For serial models, ambiguity is essential to garden-path effects; that is to say, structural ambiguity creates difficulty and confusion, i.e. larger processing difficulty; for other parallel-ranking models which are based on competition, the larger the number of fully parsed structures we have that are consistent with the partial structures seen thus far, the more competition there exists between these structures, which in turn leads to more processing difficulty. For the surprisal model, however, the more full structures which are consistent with the next word being w_i (based on the previous words and context), the larger the overall probability will be of that word indeed being w_i. The larger the probability of the next word being w_i, the easier it will be to process it. In this situation, therefore, in complete contrast to other models, when there are multiple possible parses consistent with the same upcoming word, ambiguity is predicted to actually *ease* processing, as it increases the overall probability of this word occurring. This is also consistent with a number of experimental findings in the literature (Traxler, Pickering, & Clifton Jr, 1998; Van Gompel et al., 2005, 2001).

3.2.2.2 Practically calculating predictability

Although, theoretically, including the concept of prior contextual exposure to language when generating predictions is inspired and makes logical sense, practically it can be rather hard to quantify as a measure when designing experimental stimuli and creating specific surprisal and/or entropy predictions about said stimuli. Originally, the contextual probability of a word was estimated through the measure of its *cloze probability* score. Cloze probability scoring is based on presenting participants with incomplete stimuli which are missing their final word. Researchers then count how many times participants use a specific word to complete a specific sentence. The cloze probability of a given word in a given context is then calculated as the proportion of times that particular word was used over all participants' productions, following the guidelines laid out by Taylor (1953).

Although broadly speaking there are indications that this score does indeed correlate with both psycholinguistic and neurolinguistic measures, with words with higher cloze probability showing overall shorter reaction times and smaller N400 amplitudes (DeLong et al., 2005; Wlotko & Federmeier, 2012), there are also certain problems with how researchers have used this approach. Firstly, although it is clear how cloze probability would be calculated for specific words, it is harder to conceptualise for more abstract linguistic units, such as larger syntactic structures. Furthermore, by limiting the participant to a single continuation, it is quite likely that there will be a bias for the most popular and simple word continuations. There is no indication, however, that participants do not also consider other words as likely continuations, but simply never produce them in favour of the easiest or most likely continuation. It would therefore be unrealistic to consider these non-produced but likely continuations as completely improbable, thereby maximising their predicted surprisal value. Finally, it has been argued that the nature of the task itself makes it hard to measure cloze probabilities $< 5 - 10\%$, with differences within that range considered to have no effect on measures of processing difficulty as a result (Smith & Levy, 2013), something also unrealistic.

A better way to calculate a word's probability, overcoming these shortcomings, is by taking advantage of existing large text corpora, syntactically parsed or not (depending on the aim of the study), and using them to train a probabilistic language model. There are a variety of such models, each making different (or no) assumptions about sentence comprehension mechanisms, and each with its own advantages and disadvantages when it comes to linguistic predictions. These include, for instance, n-gram or Markov models, which make no assumptions at all and do not technically take almost any context into account (i.e. no context beyond n units from the current unit); recurrent neural networks (RNN), which make more cognitive domain-general assumptions but no specifically linguistic assumptions and are either more complex, for a large number of possible words, or simpler, such as Elman's (1990), for a smaller number of possible POS; and phrase-structure grammars (PSG), which make specific linguistics assumptions regarding hierarchical tree structure creation, but which unfortunately do not appear to be the best, in fact, for linguistic predictions (Frank, Otten, Galli, & Vigliocco, 2015). These models can be trained on sentences containing the most frequently occurring words from these existing corpora and based on this training can then estimate the conditional probabilities in the above surprisal equations with great accuracy[4].

3.2.3 The Why; Uncertain Input & the Noisy Channel model

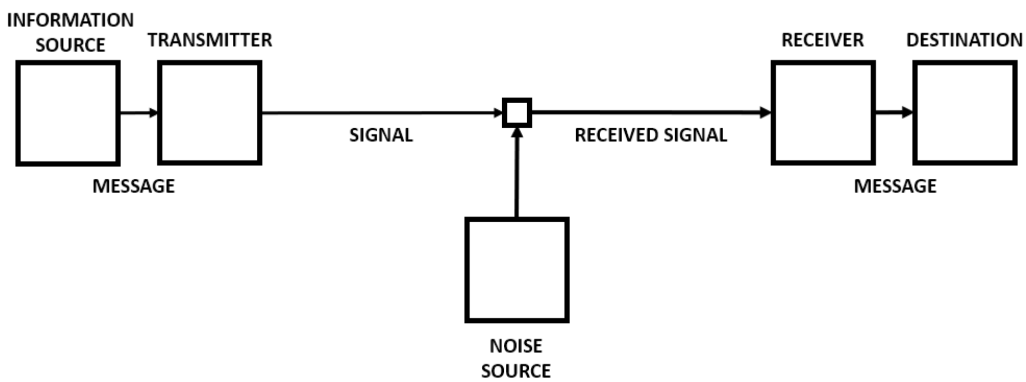
Having seen evidence that we automatically anticipate and predict upcoming linguistic material, and that this prediction is plausibly driven by the calculation and continuous update of probabilities, we now come to the question of why this might actually be the case in the first place and why we are focusing on it here.

Earlier on, we touched upon the concept of efficient communication as being the driving force behind probabilistic prediction. It stands to reason that if the parser wants to parse an incoming message as fast as possible, then it is necessary for it to calculate probabilities of upcoming material ahead of confirmatory bottom-up input. After more careful consideration, however, it also becomes clear that there is another very important factor which arguably plays a role in generating these types of predictions, and that is the fact that, unfortunately, communication is almost never perfect. Speed of communication is a perfectly adequate motivation when we assume an ideal speaker, an ideal comprehender and overall an ideal communicative situation, such as perhaps many have done in the past to create elaborate theoretical models of language. In reality, however, it quickly becomes apparent that there is no such thing as an ideal communicative situation, nor an ideal speaker or comprehender. For instance, the speaker may not always be able to find the best words to communicate their intentions; grammatical, pronunciation or spelling errors are also often made; at any point we may be surrounded by other people talking or other noises occurring which may make it hard to fully make out what the speaker is saying; we may not be paying adequate attention, either by thinking of something else, being tired, bored or simply losing our focus for a short period of time while listening, meaning we do not hear the full intended message; among many many other problems. Thus, although we enter into any communicative situation with the intent to communicate and we assume our interlocutor to do the same, quite often things do not go as smoothly as planned, primarily because there is always some form of *noise* present in our communication environment, originating from the environment itself, but also potentially from the speaker and/or comprehender.

[4]NB that for the calculation of entropy reduction, in particular, for multiple upcoming word sequences, by applying the chain rule $P(w_{t+1\ldots k}|w_{1\ldots t}) = \Pi_{i=1}^{k} P(w_{t+i}|w_{1\ldots t+i-1})$ this can quickly become computationally unwieldy as the number of unit sequences grows exponentially alongside sequence length. See Frank (2010) on how this can be done in order to maximise explanatory potential whilst at the same time reducing the chance of computational overload.

This concept was evident already to Shannon (1948) and is part of the reason his article (and subsequent book) was so incredibly influential in shaping how we perceive the communication of information. The famous mathematical depiction shown in Fig. 3.1 is seen as a diagram for general communication of any sort, although we will subsequently narrow its effects down to the field of sentence processing and linguistics for the purposes of this chapter.

Figure 3.1: Shannon 1948 General Communication Model



Specifically, this model suggests that in any form of communication, we always begin with a message that needs to be transmitted, originating from an information source and a transmitter, such as a human speaker. This speaker transmits his or her message via a signal (acoustic or otherwise) until it is received by the receiver and destination of this message, i.e. the comprehender. The comprehender must then decipher the received signal to obtain the originally intended message. Problems arise, however, when during transmission, this signal must pass through a source of potential noise which can affect it such that received and transmitted signal are not necessarily 100% identical, introducing uncertainty into the equation and making it harder to decipher the original message. Communication is deemed successful when the original message is accurately inferred through the received one, in spite of this noise.

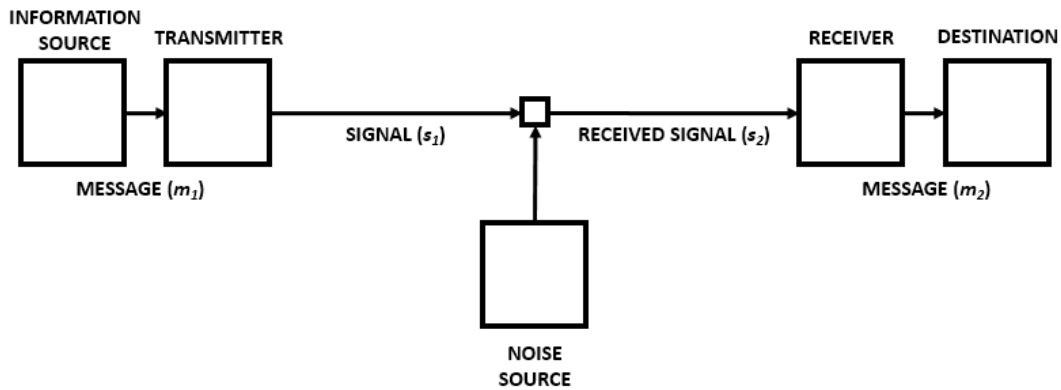
When discussing linguistic transmission of signals in particular, there are plenty of sources of noise that interlocutors might have to contend with, both literal and metaphorical, internal and external. The first, obvious source of literal external noise is, of course, the environment which communication takes place in. Background chatter in a bar, music in a nightclub, loud construction noises on a road, noise and reverberation in classrooms, all have the potential to mask and/or distort parts of the original signal (Bregman, 1994; Klatte, Lachmann, Meis, et al., 2010; Rabbitt, 1968). Other sources of 'noise', in a metaphorical sense, might still be external to the comprehender, but internal to the speaker. For instance, the speaker might make errors on any mixture of linguistic levels due to some form of speech impediment; poor command of the language as a second-language speaker; drunkenness or inattention; or even a medical condition such as Alzheimer's Disease, Broca's or Wernicke's aphasia, each of which can affect his or her disposition to grammatically frame and express complex structures and meanings; a.m.o. (Bastiaanse, Rispens, Ruigendijk, Rabadán, & Thompson, 2002; Caramazza & Zurif, 1976; Friedmann & Shapiro, 2003; Kemper, Thompson, & Marquis, 2001; Thompson, Shapiro, & Roberts, 1993). These same sources of metaphorical noise can also be internal to the comprehender, inhibiting him or her from accurately semantically dissecting linguistic messages (Almor, Kempler, MacDonald, Andersen, and Tyler (1999); Bradlow, Kraus, and Hayes (2003); Wingfield, Tun, and McCoy (2005), a.m.o.). The comprehender may also simply not be paying enough attention to the speaker or be suffering from stress. We know from various studies (Reichle, Reineberg, & Schooler, 2010; Sayette, Schooler, & Reichle, 2010; Schooler et al., 2011), for instance, that factors such as inattention, tiredness and even cigarette cravings can very easily adversely affect the quality of information being received by the comprehender (see also Schad, Nuthmann, and Engbert (2012) for possible gradation of such levels of inattention - and consequently linguistic understanding - during reading comprehension). Finally, we have much evidence that factors such as working memory capacity can also restrict our creativity with and understanding of language, and are one of the factors responsible for changes in linguistic behaviour that can become apparent with aging (Caplan and Waters (2013, 1999); Hasher and Zacks (1988); Hofmeister (2011); Hofmeister and Sag (2010); Just and Carpenter (1992), a.m.o.). Despite these nearly ubiquitous obstacles, however, it is very rarely the case that the receiver is utterly unable to reach the transmitter's intended message, or in other words that there is a complete breakdown in communication.

In 2013, Gibson, Bergen, and Piantadosi reinterpreted Shannon's original model from a purely linguistic perspective, developing an elegant, simplified model to explain how these breakdowns may be avoided in real life. Based on the annotated version in figure (3.2), to make things simpler they eschew the complications arising from ambiguity and assume that at any given time every sentence s can be 'directly and unambiguously' linked to a message m. They go on to create a simplified Bayesian probabilistic model of language comprehension shown in equation (3.3), where s_1 is the intended sentence/message on behalf of the producer and s_2 the sentence/message perceived by the comprehender.

$$P(s_1|s_2) \propto P(s_1)(Ps_1 \to s_2) \tag{3.3}$$

According to Bayes's rule, the equation in (3.3) can be read as 'The probability of the intended message (s_1), given the perceived message (s_2), is proportional to the prior probability of the intended message multiplied by the probability that this intended message

Figure 3.2: Shannon 1948 Annotated General Communication Model



could be corrupted to this perceived message'. What this means is that the parser is treated as a probabilistic decoder that generates hypotheses about upcoming and received material whilst taking into account a) the likelihood of the message itself given the sentential and extra-sentential context it is in, as well as b) the likelihood that, again, given the environment and context, this signal could be altered into the perceived one. Part of this model is essentially the same as what we described earlier when discussing surprisal. That is to say, the way we would define the prior probability of the signal is the same way that we defined surprisal, i.e. based on the frequency with which this particular linguistic unit is encountered in general and in similar sentential contexts, as well as the plausibility of the message given the sentential and extra-sentential context, i.e. the speaker and comprehender's world knowledge and shared knowledge (J. Hale (2001, 2003); J. T. Hale (2011); Levy (2008a, 2008b)). What this equation adds to the interpretation of Shannon's noisy channel model is that it takes into consideration the likelihood of signal corruption as witnessed.

This means that when calculating the posterior probability of a sentence s_1 given the perceived sentence s_2, that probability is the mathematical product of two separate, probabilistically calculated factors, each of which can co-vary with the other, but can also vary independently of the other. For instance, in a perfectly plausible and grammatical message, such as in (94a) (their (4b)), the prior probability of s_1 is very high, whilst at the same time the probability of corruption, i.e. of the intended message having been an edited version of s_1, is very small: given how likely the non-edited version of (94a) is, there is no reason to assume a different, edited version. This means that the posterior probability of s_1 can essentially be considered equal to its prior probability. In the implausible (94b) (their (4d)), on the other hand, the prior probability of s_1 is very low - it is highly unlikely that a mother would give her daughter to an inanimate object. At the same time, the number of edits required for this sentence to reach its much more plausible close neighbour in 94a is simply one, the addition of the preposition *to*, indicating that this edited sentence could potentially be the intended message, assuming some corruption. The parser is thus faced with two options; either accept the highly unlikely s_1 in (94b) and assume no changes, which leaves it with a very low posterior probability; or accept its slightly edited version in 94a whose prior probability is much higher. As long as the second option does not entail a large number of improbable edits, which would minimise the likelihood of corruption, assuming an edited s_1 would afford the parser a much higher posterior probability.

(94) a. The mother gave the candle to the daughter.
b. #The mother gave the candle the daughter.

Based on this concept of combined probabilities, Gibson et al. propose that these two factors also provide two main sources of noise that can affect the parser's predictions. On the one hand, the parser must calculate the probability of the message s_1 itself, whilst taking into account its sentential and extra-sentential context. It may be possible, therefore, to manipulate this context to make such a message more or less likely to occur. For instance, if we are in a situation where implausible or unacceptable utterances are likely, such as when we are talking to a young child with a very active imagination, someone hallucinating or otherwise confused, or someone with Wernicke's aphasia, we may be more inclined to believe that an encountered implausible utterance was indeed the intended one (or at least that there are no other close utterances which were intended). In an everyday situation, where the context does not give us any such cues, then we would be less inclined to believe an encountered implausible utterance was indeed the intended one. On the other hand, the parser must also calculate how likely various edits are to have occurred to this s_1. In situations where such edits are more likely, the parser would be more inclined to accept them, vs. a situation where they are less likely. For instance, when we are in an environment which is very 'noisy' such as at a concert or when our interlocutor has a speech impediment or does not know the language very well, we may be more inclined to believe that a perceived implausible or unacceptable structure is not in fact the intended one, but, instead, that a slightly edited version is much more likely to have been intended, but simply corrupted due to the high percentage of perceived 'noise'.

This also brings us to the question of whether there are edits which should be considered more or less likely than others to occur or whether all edits should be considered equally likely, both in terms of number and type of edits. Logically speaking, one would assume that the probability of s_1 being altered to s_2 would increase as the number of edits required for the two to coincide decreases, i.e. a smaller number of edits ((94a)→(94b): 1 deletion) should always be more likely to occur than a larger number of edits ((95a→(95b): 2 deletions), all other things being equal. This also makes sense based on the general Bayesian size principle (Tenenbaum, 1999), the principle that the smaller hypothesis will always have a greater probability than the larger hypothesis.

(95) a. The ball **was** kicked **by** the girl.
b. #The ball kicked the girl.

In terms of the type of edits available, Gibson et al. assume that there are only two types of string edits: insertions and deletions. Previous noisy channel model iterations, such as in (Levy, 2008a, 2008b), make a similar assumption and treat both these edits as independent and equal in terms of probability, assuming a symmetric Levenshtein distance[5] noise model. Gibson et al., however, differentiate between the two, predicting that deletions should be more likely to occur than insertions. They argue that this prediction follows from the Bayesian size principle, which we just saw (Tenenbaum, 1999). In this situation, it is more plausible that a specific word may have been accidentally deleted from the utterance during message transmission, rather than that a specific word may have been erroneously inserted. The former requires an existing word to be deleted from the limited area of a sentence, whereas the latter requires a spurious word to be selected from the transmitter's entire lexicon and randomly inserted into the utterance. Out of the two, therefore, they argue that the former represents the smaller hypothesis and, as such, should be afforded greater probability.

Based on the above predictions, Gibson et al. conducted three web-based studies in which they measured how often participants interpreted an implausible sentence as a) its close, plausible neighbour sentence or b) literally as presented. In all three experiments they varied the number of edits required to get from the plausible sentence to the implausible one (1 or 2) as well as the type of edits required (insertion(s) vs. deletion(s) vs. one of each). From experiment to experiment, they also varied the amount and type of 'noise' present in the experimental context, either simulating an environment where corruption was more likely or one where intended implausible utterances were more likely. This was done by manipulating the base rate of a) syntactic errors in the fillers (thereby increasing the prior contextual probability of corruption); and b) implausible sentences in the fillers (thereby increasing the prior contextual probability of an implausible utterance). They found that, overall, their predictions were borne out. In terms of the type of edits, comprehenders appeared more likely to accept a structure where a specific element has been spuriously deleted, rather than one where a specific element has been incorrectly added in. In terms of the sheer number of changes, they were also more likely to accommodate for an utterance with only one necessary change as opposed to two or more, with this acceptance rate decreasing

[5] Levenshtein distance is a metric of the difference between two strings, with its value indicating the number of single character edits required to get from the source string to the target string (Levenshtein, 1966).

with each additional change required. Perhaps more interestingly, however, they also found a significant effect of the factor of contextual 'noise' from experiment to experiment, i.e. they found that a) participants were significantly more likely to interpret an implausible utterance literally when there was more 'implausible noise' in the stimuli in the form of implausible fillers, compared to when there was 'syntactic error noise' or no 'noise'; and b) participants were significantly more likely to interpret an implausible utterance as its plausible neighbour when there was more 'syntactic error noise' in the fillers compared to when there was 'implausible noise' or no 'noise'.

With these manipulations, Gibson et al. provide evidence for the parser indeed appearing to work as a Bayesian probabilistic inference machine operating through a noisy channel, taking into consideration not only the semantic probability of upcoming material, but also the contextual noise it appears in.

One final point which should be made concerns how exactly prior expectations may affect the parser's decision-making beyond what we have described above. Although this is not strictly considered in Gibson et al. (2013), in Levy (2008a) another prediction for a probabilistically-driven parser would be not only in terms of plausibility of one sentence vs another, but also in terms of sheer likelihood, i.e. frequency, of one sentence structure vs another. One could predict that even in cases where we are dealing with two very close, in terms of Levenshtein distance, utterances, *both* of which are entirely acceptable and plausible, prior expectations in terms of frequency of occurrence could plausibly lend more probability to the message in the *edited* version; that is to say, if the edited version is much more frequently occurring, this could push the parser to accept it over the unedited one, despite the latter also being perfectly acceptable.

3.3 Sluicing as a Noisy Channel

3.3.1 'P-Stranding' under Sluicing

3.3.1.1 Cross-linguistic Predictions

In the remainder of this thesis, I will be entertaining a noisy channel hypothesis for why we may be seeing people accept P-stranding under sluicing in languages which are both overtly non-P-stranding, but also do not allow an alternative, acceptable structure at the *e*-site, such as in Greek and German.

As mentioned in Chapter 1, one of the reasons that sluicing and ellipsis, more generally, have garnered so much attention over the past few decades is that they are one of the most profound examples of linguistic meaning being conveyed through silence. This lack of overt evidence is also the reason why analysing ellipsis has proven so difficult, controversial and almost impossible to easily convince readers of one theoretical analysis over another. I will argue here that this lack of overt structure is also an example of a *natural noisy channel*. In having to subconsciously fill in the missing structure either by recalling antecedent structure through a pointer mechanism or by rebuilding the structure in-situ based on the available overt material, we automatically open the door to all sorts of potential 'noise' based on our less-than-perfect working memory compared to, say, when the same structure is overtly expressed. In order to properly interpret ellipsis, we must rely on these imperfect memory mechanisms in conjunction with our narrower and wider contextual knowledge and expectations. This reliance on previously seen material is possibly one of the reasons why we find all sorts of instances of mismatching under ellipsis being entirely or partially accepted, such as voice mismatching under VP-ellipsis ((96), (97), Dalrymple, Shieber, and Pereira (1991); Kehler (2000); see also Arregui, Clifton, Frazier, and Moulton (2006); Johnson (2001); Kim, Kobele, Runner, and Hale (2011), though cf. Sag (1976) and Williams (1977) a.o. for counter-examples), or number- and gender-mismatching in sluicing (see also Chapter 1) and NP-ellipsis ((98), (99), see Merchant (2014) for Greek; Eguren (2010); Leonetti (1999); Masullo and Depiante (2004) for Spanish), all alongside the main tendency for parallelism between antecedent and *e*-site.

To put it differently, one could argue that if our previous expectations can not only affect how difficult it is to parse input, but can also shape our perception of the input itself, then a less-than-complete input string is likely to be even more liable to being shaped by such prior expectations, with us leaning on these expectations and knowledge to help us fill in the gap. In other words, if plausibility and frequency can make us interpret overt utterances differently to the way they are being literally presented, then surely a lack of overt material makes it even more likely that our interpretation of elided material will be affected by these expectations.

(96) Dalrymple et al. (1991)
In March, four fireworks manufacturers asked that the decision be reversed, and on Monday the ICC did [~~reverse the decision~~].

(97) Kehler (2000)
This problem was to have been looked into, but obviously nobody did [~~look into this problem~~].

(98) Merchant (2014), Gender mismatching in Greek

O	Petros	ine	kalos	daskalos,	ala	i	Maria
*Det.*M.NOM	*Peter.*M.NOM	*is.*3SG	*good.*M.NOM	*teacher.*M.NOM	*but*	*Det.*F.NOM	*Maria.*F.NOM

ine	mia	kakia.
*is.*3SG	*a.*F.NOM	*bad.*F.NOM

'Petros is a good teacher, but Maria is a bad one.'

(99) Merchant (2014), Number mismatching in Greek

O jiatros ine daskalos, ala i dikijori
*Det.*M.NOM *doctor.*M.NOM *is.*3SG *teacher.*M.NOM *but Det.*M.NOM.PL. *lawyers.*M.NOM.PL.
ochi.
no

'A doctor is a teacher, but lawyers are not.'

By treating sluicing in this way as a potential noisy channel, through which the parser must calculate the posterior probability of encountered material, this also allows us to explore how the string edits proposed by Gibson et al. (2013) and Levy (2008a, 2008b) might appear and interact with the prior likelihoods of plausible sentences in generating these posterior likelihoods. I will argue, more specifically, that what looks like P-stranding under sluicing in languages which have no alternative *e*-site structure is in fact P-pied-piping with a deleted P, i.e. an example of string deletion through a noisy channel. Based on the Bayesian equation in (3.3) (repeated below), when faced with the P-less utterance in (100a), the parser has two options: either a) accept this unacceptable string literally as presented, which, being unacceptable, has a very low prior likelihood; or b) assume that this is not the intended string and that the very close string in (100b) was, instead, intended. In option (a), the posterior likelihood of (100a) is equal to the very small prior likelihood of (100a) multiplied by the high likelihood of no string edits; in option (b), the posterior likelihood of (100b) is equal to the very high prior likelihood of (100b) as an utterance, multiplied by the likelihood of 1 string edit, in itself slightly lower than the likelihood of having no string edits at all, but still relatively high, according to logic and Gibson et al.'s findings. Thus, when we compare the two posterior likelihoods these options represent, thanks to the high prior likelihood of (100b), the posterior likelihood in option (b) should be much higher than the posterior likelihood in option (a); this is under the proviso that a single deletion is not a completely improbable edit, an assumption which, given the results in Gibson et al. (2013), would appear to be correct. Hence, the parser should theoretically be more inclined to select option (b), i.e. to infer that the intended message is in fact the one in (100b) and not the literally presented one in (100a).

$$P(s_1|s_2) \propto P(s_1)(Ps_1 \rightarrow s_2) \tag{3.3}$$

(100) a. **P-stranding under sluicing**

I Maria chorepse me kapjon, alla de thimame
*Det.*F.NOM *Maria.*F.NOM *danced.*3SG *with someone.*M.ACC *but* NEG *remember.*1SG

pjon.
who.M.ACC

'Maria danced with someone, but I don't remember who.'

b. **P-pied-piping under sluicing**

I Maria chorepse me kapjon, alla de thimame
Det.F.NOM *Maria*.F.NOM *danced*.3SG *with someone*.M.ACC *but* NEG *remember*.1SG
me pjon.
with *who*.M.ACC

'Maria danced with someone, but I don't remember **with** who.'

(101) a. **P-stranding overtly**

*I Maria chorepse me kapjon, alla de thimame
Det.F.NOM *Maria*.F.NOM *danced*.3SG *with someone*.M.ACC *but* NEG *remember*.1SG
pjon chorepse me.
who.M.ACC *danced*.3SG *with*

'Maria danced with someone, but I don't remember who she danced with.'

b. **P-pied-piping overtly**

I Maria chorepse me kapjon, alla de thimame
Det.F.NOM *Maria*.F.NOM *danced*.3SG *with someone*.M.ACC *but* NEG *remember*.1SG
me pjon chorepse.
with who.M.ACC *danced*.3SG

'Maria danced with someone, but I don't remember with who(m) she danced.'

At this point, I should point out that this process does not mean that P-stranding under sluicing in languages such as Greek should be as acceptable and easy to parse as P-pied-piping nor that it should be considered the norm, i.e. that we should completely overturn the P-Stranding Generalisation. Instead, I argue that, technically, pied-piping the P should still be considered the rule and the easiest structure to parse of the two, with P-stranding being the exception. That is to say, in selecting option (b) above and inferring a string-edited version of the encountered utterance, there could potentially be an associated *processing cost* compared to if, say, the parser were to simply encounter the P-pied-piped structure in (100b) as is, with no string edits being necessary. If this is true, then, all other independent processing pressures aside, this processing cost would be equal to the cost of inferring the string-edited plausible sentence in (100b) and, if everything we have seen in the above sections is correct, should theoretically be equal to the difference in posterior likelihoods between having to infer (100b) from (100a) vs. having (100b) to begin with. To phrase this in the terminology of surprisal, the difference in posterior likelihoods between these two sluices, one with an overt P, one without, would translate to a difference in the surprisal associated with each, which in turn would translate to a difference in cognitive effort required to parse each, such that the sluice without an overt P would be harder to parse compared to the one with an overt P. This difference in processing difficulty could, furthermore, translate to differences in acceptability, as has been argued by others before (Hofmeister et al., 2013), such that we would anticipate apparent P-stranding under sluicing to be rated lower on the acceptability scale compared to P-pied-piping.

At this point, it may be worth theoretically comparing this P-deletion under sluicing to P-deletion overtly. If our previous assumption that sluicing serves as a natural noisy channel is correct, then we would anticipate that people would be more likely to accept P-deletion under sluicing compared to overtly, given the former is in the context of a noisy channel. In other words, if we were to compare the two, we would anticipate that P-deletion overtly is rated lower on the acceptability scale than P-deletion under sluicing, which, in turn, should be lower than P-pied-piping under sluicing, as concluded above.

Moving forward with this hypothesis, if we are willing to assume that sluicing can and does provide a 'noisy' medium as described above, then this phenomenon should also be true cross-linguistically. If so, a question which naturally arises at this point is how would this noisy channel model of sluicing fare with respect to other overtly non-P-stranding languages which, however, appear to allow for an alternative structure at the *e*-site, such as a copular or resumptive source? The simple answer to this question is that when faced with a potential P-stranding structure under sluicing in these languages, thanks to the available acceptable alternative structure, the parser need not engage in the above weighing of probabilities at all. To use a simple example from Brazilian Portuguese (based on the judgements in Rodrigues et al. (2009)), given the potential structures in (102), when the parser is faced with (102b) it need not bother considering (102a) at all. Instead, it is presented with a very simple solution: to assume the acceptable alternative structure in (102d), which is perfectly plausible as is and requires no string edits. In other words, the parser should automatically select the option with the highest posterior likelihood out of all those available. Moreover, given this structure is acceptable and need not be edited, there should be no processing cost attached to it. The acceptability of (102b) should, therefore, be equal to the acceptability of (102d). Again, assuming the cognitive effort expended to parse a structure is only affected by these probabilities, the difference in processing cost (and thereby acceptability) between (102b) and (102a), furthermore, should be equal solely to the difference between the prior likelihoods of each. That is to say, if pied-piping under sluicing is as plausible as this alternative structure, then there should be no difference between the two, either in processing cost or acceptability. What happens, however, if the alternative structure in

102d is somehow blocked? In that case, the parser would be forced to engage in the same weighing of probabilities as in other non-P-stranding languages such as Greek and German, i.e. as though no alternative were available.

(102) A Maria dançou com alguém, mas . . .
the Maria danced with someone but
'Maria danced with someone, but . . .'

a. . . . eu não sei com quem.
. . . I not know with who
'. . . I don't know with who.'

b. . . . eu não sei quem.
. . . I not know who
'. . . I don't know who.'

c. * . . . eu não sei quem a Maria dançou.
. . . I not know who the Maria danced
'. . . I don't know who Maria danced with.'

d. . . . eu não sei quem é com a qual a Maria dançou.
. . . I not know who is with the that the Maria danced
'. . . I don't know who it is with which she danced.'

This line of reasoning predicts two distinct sets of results not only between different overtly non-P-stranding languages, but also within the same language, dependent upon the availability or not of an alternative *e*-site structure (see table 3.1 for predictions). Specifically, in situations where an alternative, acceptable structure is available, what appears as P-stranding under sluicing (row *f*) should have the same acceptability as this alternative structure does overtly (row *c*) (e.g. languages L1 & L2). If this alternative is as plausible as P-pied-piping overtly (row *a*) (e.g. L1), then we should find no deterioration of acceptability under sluicing for what looks like P-stranding (row *f*) when compared to P-pied-piping (row *d*), given the alternative, plausible structure can be assumed. If this alternative is slightly less plausible than P-pied-piping overtly (e.g. L2, row *c* vs. row *a*), then we should find this small difference in acceptability translate to such a difference under sluicing as well (row *f* vs row *d*). This minor difference between the two aside, in both L1 and L2, this 'P-stranding' structure should be much more acceptable than overt P-stranding (row *b*). If, however, this alternative structure is blocked or the language does not have such a structure to begin with (e.g. L3), then the parser must adapt and engage in the weighing of probabilities as described above. If the difference between the plausible P-pied-piped sluice and the implausible P-stranded sluice is simply the deletion of a P, then the editing process for reaching one structure from the other should be relatively straightforward. As a result of the processing cost associated with accepting the more plausible, but string-edited, P-pied-piped version, P-stranding under sluicing in this situation (row *e*) should be considerably more acceptable than overtly, but not as acceptable as a) P-pied-piping under sluicing (row *d*) or b) P-stranding under sluicing when an alternative *e*-site structure is available (row *f*), given neither of these requires any string edits.[6][7].

Table 3.1: Cross-linguistic P-Stranding Predictions for Non-P-Stranding Languages

'The Maria danced with someone, but not know ...			Acceptability*			
	Overt Structure	*e*-site	L1	L2	L3	L4
Overt	a) ... *with who she danced.'*	N/A	6/7	6/7	6/7	6/7
	b) ... *who she danced with.'*	N/A	2/7	2/7	2/7	6/7
	c) ... *who it was with which she danced.'*	N/A	6/7	5/7	2/7	6/7
Sluicing	d) ... *with who*	[she danced].'	6/7	6/7	6/7	6/7
	e) ... *who*	[she danced with].'	4/7	4/7	4/7	6/7
	f) ... *who*	[it was with which she danced].'	6/7	5/7	N/A	6/7

* *Using a scale of 1 - 7, with 1 being completely unacceptable and 7 being completely acceptable.*

For the sake of completeness, it should be pointed out here that following the same reasoning, an overtly P-stranding language such as English should show no deterioration of P-stranding under sluicing compared to a) P-stranding overtly or b) P-pied-piping under sluicing (L4). However, there are no existing studies to my knowledge that actually test this (although see Chapter 5 for future directions). The difference in acceptability between the two structures under sluicing would be tied to the frequency with which embedded interrogative clauses occur with P-pied-piping vs. with P-stranding, with higher frequency likely leading to a higher acceptability rating; for instance, in English in particular, it has been argued that P-stranding is generally preferred over P-pied-piping (Cable and Harris (2011); Heck (2008), see also Chapter 4 here), although, once again, adequate testing is necessary to confirm this.

[6] This set of acceptability predictions comes with the caveat that we are not considering the effects of any other, independent processing costs here; the purpose of this table is simply to illustrate the predicted processing costs associated with inferring the posterior probability of a message under sluicing, all other things being equal.

[7] As is evident, we are assuming here that differences in processing costs will translate to differences in acceptability, as Gibson et al. (2013) do as well. Although the relationship between these two behavioural measures may not be precisely linear, we are assuming that overall the directionality of behaviour is preserved. This may be an inaccurate assumption and perhaps one that not everyone agrees with. Unfortunately, there have not been enough studies on this matter to date to make a clear decision one way or the other, though see Hofmeister et al. (2014, 2013); Staum Casasanto, Hofmeister, and Sag (2010) for arguments in favour. However, even though I accept that measured processing costs may not always translate into measurable acceptability effects, I believe it reasonable to make a prediction regarding the opposite, i.e. that if we find a measurable difference in acceptability ratings, then it is possible that this difference may be due to measurable differences in the cost of parsing each structure, subject to confirmatory further testing.

As a final side-note, one effect which could perhaps be predicted cross-linguistically and that is not included in the table in order to keep things as simple as possible, is that when no string edits are necessary, sluicing may be *more* acceptable than overt continuations, given the repeated name penalty found in other processing studies (Almor, 1999; Garrod et al., 1994; Gordon et al., 1993, 1999; Swaab et al., 2004)

3.3.1.2 Cross-linguistic Data

Now that we have a clearly defined set of predictions that this hypothesis makes, it is time to see whether they capture the actual acceptability data presented in the previous chapter. Specifically, given Greek and German do not have acceptable alternative structures available for P-stranding, as argued in Chapter 2, this places them in the 'L3' category in Table 3.1 above. In terms of acceptability, this means that although P-stranding overtly should be unacceptable, it should be significantly more acceptable under sluicing, albeit significantly less so than either P-pied-piping under sluicing or overtly, with the latter two being equal to each other (assuming no other processing pressures are concurrently having an effect, such as, for instance, the repeated name penalty). Figure 3.3 below presents the case-matching data from Experiments 1 (Greek) and 4 (German).

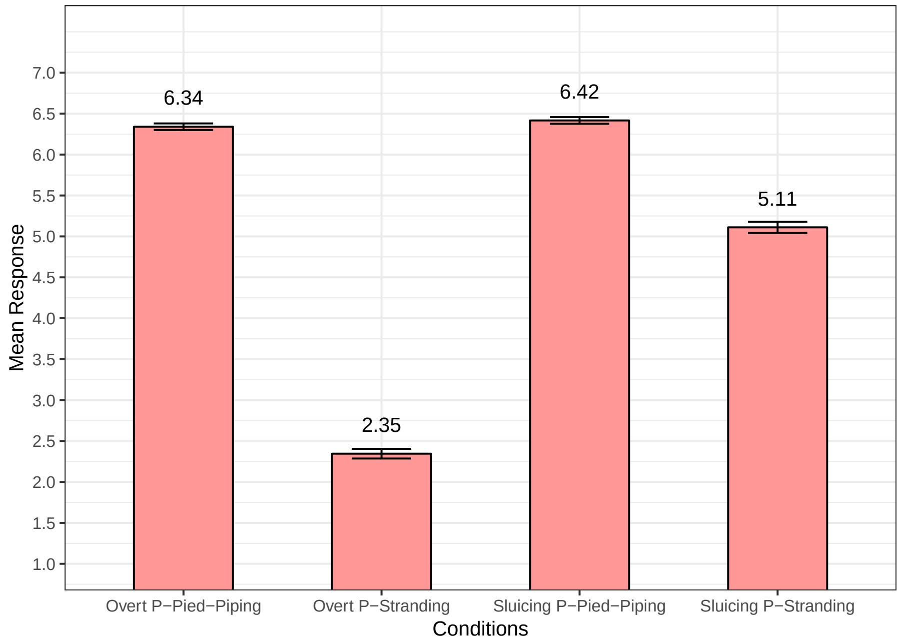
As seen from an examination of these data sets, this predicted pattern is exactly what was found. To reiterate the analysis results for each study, there was a significant interaction found between the factors of P-stranding and Sluicing such that P-stranding was significantly more acceptable under sluicing than it was overtly, with no such difference for P-pied-piping. There was also a significant main effect of P-stranding with P-pied-piping conditions always being significantly better than P-stranding ones, whether overtly or under sluicing. Furthermore, as also mentioned in the previous chapter, although there are theories which can explain apparent P-stranding under sluicing in the absence of alternative structures, these do not capture the acceptability gradient found in both studies, with P-pied-piping being significantly better than P-stranding under sluicing, making this the only theory to capture this pattern.

As seen in section 2.5.3.3, an additional analysis of the combined datasets including the between-subjects factor of 'Language' found a significant four-way interaction between the factors of Language, Case-matching, P-stranding and Sluicing, with P-stranding with case-matching being significantly better under sluicing for Greek compared to German. Bearing in mind the differences between studies (such as different subjects and items, as well as recruitment method - one was through Facebook friends and acquaintances without reimbursement per subject, the other was through a survey website (Prolific), with reimbursement per subject), which may have affected scoring, and that the German data have not been replicated yet, there are a few possible explanations for this difference within a noisy channel approach. One possibility is that the frequency with which P + NP.ACC occurs as a

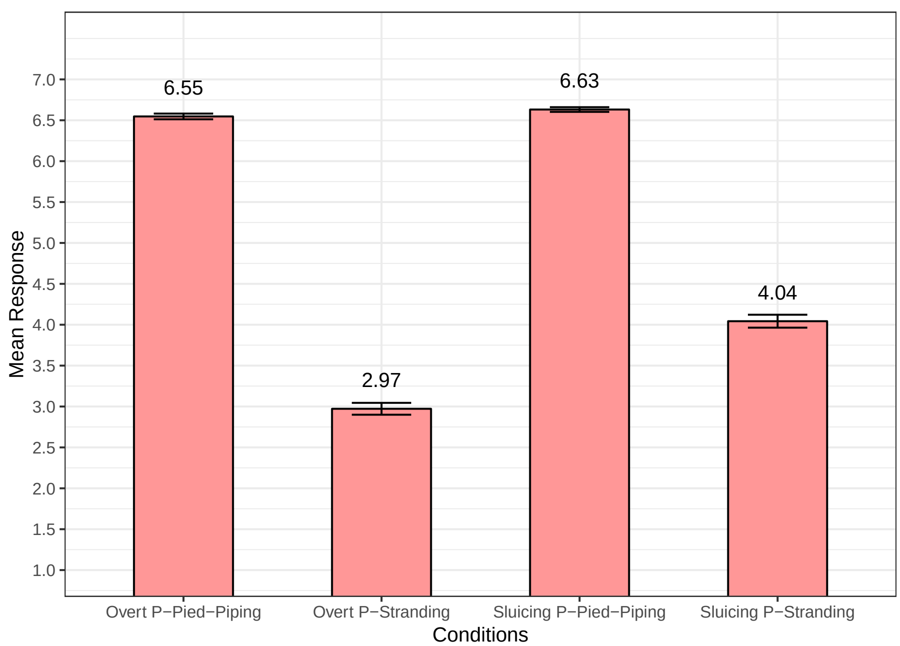
Figure 3.3: Greek & German Case-Matching Results

Explanatory Notes: Each figure depicts Mean response per condition with SEM error bars

(a) Greek Mean Response Barplots (Experiment 1)



(b) German Mean Response Barplots (Experiment 4)



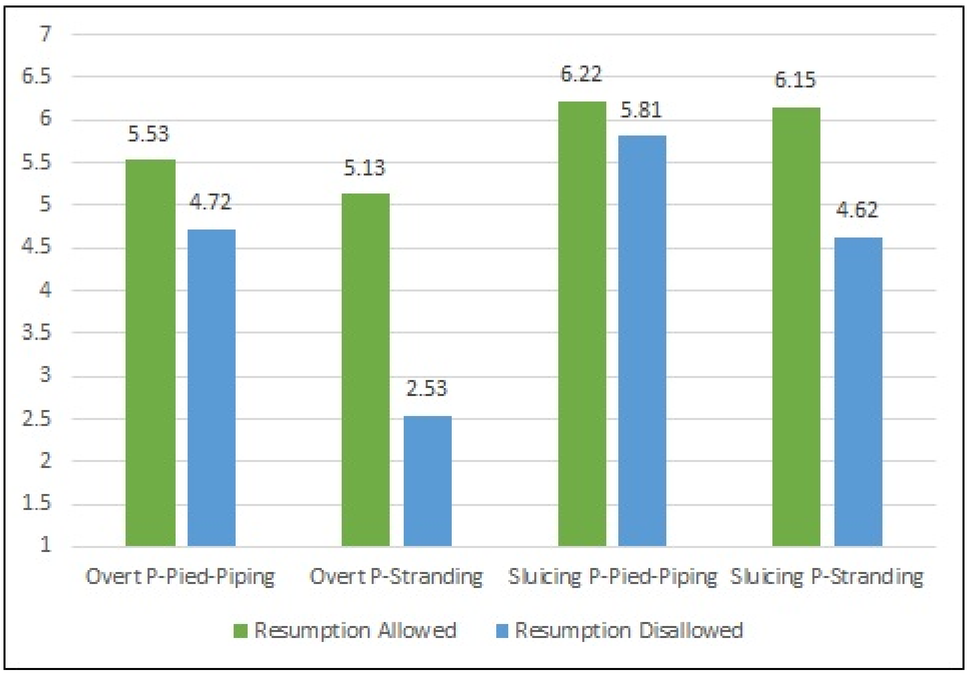
verb complement in Greek is larger than the frequency with which P + NP.DAT occurs as a verb complement in German. If this is true, then this higher frequency would translate to an increased prior probability of occurring; this in turn would increase the structure's posterior probability, thus making it easier to parse. The fact that all Ps predominantly occur with Accusative-marked NPs in Greek, whereas Ps are split between Accusative-, Dative- and even Genitive-marked NPs in German, would also increase the frequency of occurence of P + NP.ACC compared to P + NP.DAT in German. This is something that a corpus analysis would be able to shed more light on.

With respect to the other languages seen in Chapter 1, for the largest part these were not adequately tested to show whether the above predictions hold true. The exception to this is Saudi Arabic (SA), an overtly non-P-stranding language which does sometimes allow for an alternative *e*-site structure, and which was tested by AlShaalan and Abels in 2019 following Molimpakis (2016) and the guidelines presented here in Chapter 2. Based on Table 3.1 above, SA should behave as an L1 or L2 in those situations where an alternative, resumptive structure is available (with the difference between L1 and L2 being only how acceptable the alternative overt structure is compared to overt P-pied-piping). In other words, when resumption is an option, P-stranding under sluicing should be rated as acceptably as this alternative source is rated overtly (once again, assuming no other processing pressure differences between the two). On the other hand, when resumption is blocked, i.e. there is no alternative structure available, then SA should behave similarly to Greek and German, i.e. as belonging to category L3. That is to say, P-stranding under sluicing should now be less acceptable than when an alternative source is available, but still significantly more acceptable than it is overtly. This is exactly what AlShaalan and Abels found, with the relevant results (from their experiment 4, in particular) presented in Figure 3.4. Specifically, they found that when resumption is allowed, P-stranding and P-pied-piping were equally acceptable overtly, but also under sluicing. When resumption was blocked, P-pied-piping was acceptable both overtly and under sluicing; P-stranding, on the other hand, was unacceptable overtly, but significantly more acceptable under sluicing. That having been said, even under sluicing it was significantly less acceptable than a) P-pied-piping under sluicing when resumption was disallowed; b) both P-stranding and P-pied-piping under sluicing when resumption was allowed[8]. Furthermore, although an approach based only on the availability of syntactic structure at the *e*-site, such as taken by AlShaalan and Abels, does not explain this second pattern of behaviour, i.e. how P-stranding can be acceptable when the alternative structure is blocked, a noisy channel interpretation of the data as provided here does successfully capture it.

[8]As a side note, with respect to the differences shown here between sluicing and overt continuations, it appears as though something along the lines of the repeated name penalty is having an effect on the acceptability of sluicing in SA, with all sluicing conditions being rated on average as more acceptable than their overt counterparts. This is not something I found in Greek or German, so it is interesting to observe this pattern here, making this cross-linguistic difference something perhaps worth looking into more.

Figure 3.4: AlShaalan and Abels (2019) Mean Acceptability Ratings Barplot

Explanatory Notes: Results from AlShaalan and Abels (2019), Experiment 4, re-ordered.



In conclusion, as far as P-stranding is concerned, a noisy channel hypothesis as laid out above perfectly captures the interesting cross-linguistic and intra-linguistic acceptability patterns observed in all sluicing acceptability data collected thus far. Furthermore, I would predict that all overtly non-P-stranding languages should behave similarly to either Greek and German or SA, depending on whether an acceptable alternative structure is available for P-stranding under sluicing. For overtly P-stranding languages, as mentioned above, these should find both P-stranding and P-pied-piping acceptable, both overtly and under sluicing.

3.3.2 Case-Mismatching under Sluicing

Although the focus of this dissertation concerns P-stranding under sluicing, in the interest of completeness, we will also look here through the lens of a noisy channel model at the other two main features of sluicing, namely the case-matching generalisation and the fact that sluicing appears to ameliorate most islands.

Let us look first at the second form-identity generalisation put forward in Merchant (2001), together with the Fit Amendment (Abels, 2016), namely that in case-marking languages, the sluice remnant must appear in the case it would normally be assigned had it been followed by an overt copy of the antecedent, save for the *wh*-phrase. Under a noisy channel hypothesis, predictions regarding how the parser might deal with case mismatching of this remnant under sluicing are, unfortunately, not as straightforward as those for P-stranding, given case mismatching does not directly map on to the types of string edits considered in the literature, i.e. insertion or deletion. When one examines instances of case mismatching, such as in (103a), the more accurate description of the difference between this and case-matching in ((103b), differences highlighted in bold) would appear to be one of *substitution*. Although substitution is not considered in Levy (2008a, 2008b) or Gibson et al. (2013), in order to be able to make direct predictions based on the results of the latter, as we did for P-stranding, one could potentially reframe substitution as a form of deletion, followed by insertion (or vice versa). The more straightforward forms of deletion and insertion here would involve the whole word, as in (103c), however it is not impossible to hypothesise that these string edits might also occur *within* the word itself, depending on the model's granularity (e.g. *pjo-* ~~*n*~~ *-s*) [9]. For now, given we are dealing with the deletion of a whole word when considering the case of P-stranding under sluicing, in the interest of consistency, let us maintain the same granularity in our predictions for case-matching and assume that it is an entire word which is being deleted and inserted, although I would tentatively assume that the same predictions should theoretically also hold for smaller (or

[9]This extension is not considered by Gibson et al. (2013), however everything else being equal, it would appear to follow theoretically from J. Hale (2001, 2003); Levy (2008a, 2008b); Shannon (1948) as well as practically from applications of a noisy channel to corrections of orthographic errors which also concern only part of a word. That having been said, it is also possible that string edits within the same word are subject to different or additional processing costs or rules.

larger) grain sizes.

(103) I Maria filise kapjon, alla de thimame...
Det.F.NOM *Maria*.F.NOM *kissed*.3SG *someone*.M.ACC *but* NEG *remember*.1SG..
'Maria kissed someone, but I don't remember...'

a. ...pjo**s**.
...*who*.M.NOM
'...who.'

b. ...pjo**n**.
...*who*.M.ACC
'...who(m).'

c. ...[~~pjon~~] pjo**s**
...*who*.M.ACC *who*.M.NOM

d. ...pjo**s** filise.
...*who*.M.NOM *kissed*.3SG
'...who she kissed.'

e. ...pjo**n** filise.
...*who*.M.ACC *kissed*.3SG
'...who(m) she kissed.'

(104) I Maria chorepse me kapjon, alla de thimame...
Det.F.NOM *Maria*.F.NOM *danced*.3SG *with someone*.M.ACC *but* NEG *remember*.1SG..
'Maria danced with someone, but I don't remember...'

a. ...pjo**n**.
...*who*.M.ACC
'...who.'

b. ...me pjo**n**.
...*with who*.M.ACC
'...with who.'

c. ...pjo**n** chorepse me.
...*who*.M.ACC *danced*.3SG *with*
'...who(m) she danced with.'

d. ...me pjo**n** chorepse.
...*with who*.M.ACC *danced*.3SG
'...with who(m) she danced.'

$$P(s_1|s_2) \propto P(s_1)(Ps_1 \rightarrow s_2) \tag{3.3}$$

Based on the equation in 3.3 (repeated above), the Bayesian size principle and Gibson et al.'s results, if 'correction' of case mismatching involves two string edits and that of P-stranding involves only one, then - assuming the same prior probability for both structures - we would anticipate the posterior probability of a case-mismatched sluice (103b) to be smaller than that of a P-stranded sluice (104a). If, furthermore, we continue to assume that the processing cost of input is inversely correlated with its posterior probability, then we would assume that it is harder to parse case mismatching under sluicing compared to P-stranding. Based on this, we might predict that case mismatching would show lower acceptability ratings under sluicing than P-stranding does. It is unclear whether the processing cost of these two string edits when added together would translate to acceptability ratings which are still higher than those for overt case mismatching (103d), as in the case, for instance, of P-stranding overtly and under sluicing, or not. It is still possible that we would find some small effect of a noisy channel, such that case mismatching is also more acceptable under sluicing compared to overtly, however this would probably be to a smaller degree than we would expect for P-stranding.

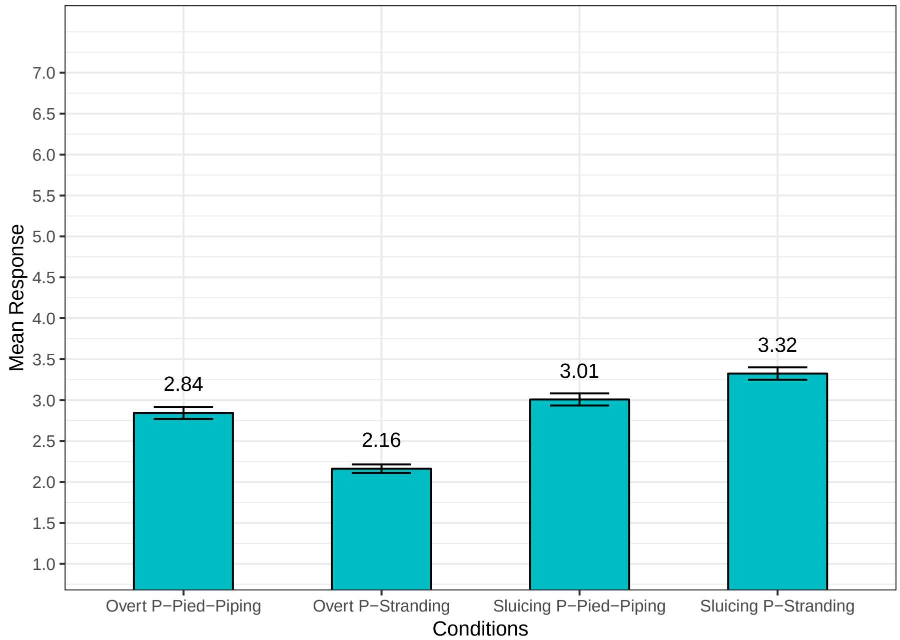
Bearing these predictions in mind, let us now look at the case-mismatching results from Greek and German, presented in Figure 3.5, as well as how they compare to the case-matching, P-stranding ones examined above. In both experiments, results showed that case-mismatching was unacceptable overtly, more so when combined with P-stranding compared to P-pied-piping, the two overt violations appearing to have an additive effect on acceptability ratings. The results are more interesting when it comes to sluicing, however. Specifically, in both languages, there was a significant interaction of Case-matching, Sluicing and P-stranding, such that although case-mismatching had the worst ratings with P-stranding overtly, as just mentioned, it had the highest ratings with P-stranding under sluicing. That is to say, even though case-mismatching with P-stranding requires more string edits compared to case-mismatching with P-pied-piping in order for both to reach case-matching P-pied-piping under sluicing (i.e. 3 string edits in (105e): 2 deletions, 1 insertion; vs. 2 string edits in (105f): 1 deletion, 1 insertion), in both languages the latter was rated as less acceptable than the former. One quite likely reason for this is that case-mismatching with P-pied-piping as in (105f) contains an overt reminder of the remnant's incorrect case directly adjacent to it in the form of the overt P. That is to say, regardless of any potential continuations for this sluice - which may or may not be argued to contain a feature-marking constituent requiring the remnant to appear in a specific case - the remnant here appears most obviously as the complement of this P. Given P complements are never Nominative-marked in either language (as mentioned in the previous section (3.3.1.2)), this would make the remnant's case-marking particularly jarring, all other things aside. With P-stranding, on the other hand, there is no such overt reminder (105e).

There are two further interesting findings with respect to this condition of P-stranding

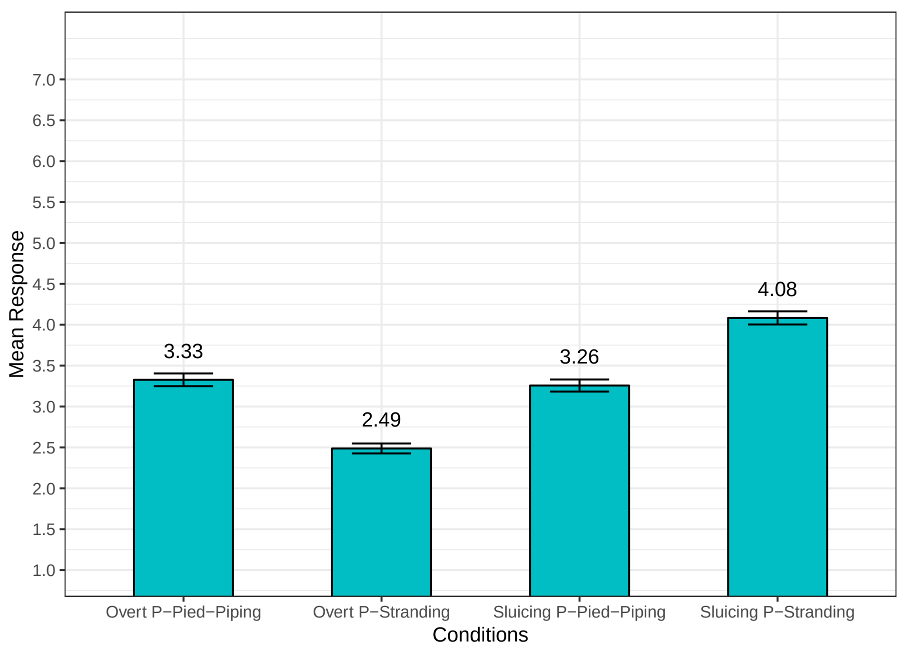
Figure 3.5: Greek & German Case-Mismatching Results

Explanatory Notes: Mean response per condition with SEM error bars.

(a) Greek Mean Response Barplots (Experiment 1)



(b) German Mean Response Barplots (Experiment 4)



with case-mismatching under sluicing. In both languages, as mentioned above, this condition is rated as more acceptable than P-pied-piping with case-mismatching (Greek Sluicing P-Stranding Case-Mismatching Mean: 3.324; Greek Sluicing P-Pied-Piping Case-Mismatching Mean: 3.008; Mean Difference: 0.317; German Sluicing P-stranding Case-Mismatching Mean: 4.083 ; German Sluicing P-Pied-Piping Case-Mismatching Mean: 3.257 ; Mean Difference: 0.827). In Greek, however, this condition is significantly less acceptable than P-stranding with case-matching (Mean: 5.111; Mean Difference: 1.786); in German, on the other hand, the two are rated as equally acceptable (Mean: 4.043; Mean Difference: 0.040). Given in both languages the two conditions are the same numbers of string edits away from an acceptable, case-matching P-pied-piping condition, how can this cross-linguistic difference be explained via a noisy channel hypothesis? One possibility here is that, when faced with both case-mismatching and P-stranding, the parser does not in fact entertain a case-matching P-pied-piping continuation as the likely intended utterance, but instead considers one that is fewer string edits away. Specifically, using the Greek example in (105), with German being very similar, P-stranding with case-mismatching (105e) is theoretically 3 string edits away from P-pied-piping with case-matching (105b). On the other hand, a pseudosluice-style continuation, as in (105i), is 0 string edits away. Therefore, in terms of sheer number of string edits, a pseudosluicing continuation like this is a much more likely intended structure out of the two. In order to calculate the highest posterior probability out of the two, this leaves the parser with the task of weighing up the prior probability of each of these structures given the context, i.e. a sluice continuation featuring P-pied-piping with case-matching vs. one with pseudosluicing. This prior likelihood is linked to the frequency with which different types of overt embedded questions targeting a PP-contained indefinite occur. When targeting such a PP indefinite in the antecedent, if P-less cleft-like embedded questions are more frequently occuring than ones containing *wh*-movement with P-pied-piping, then this would translate to a pseudosluice-style continuation here having a higher prior likelihood than one with P-pied-piping and case-matching, which in turn would translate to a higher posterior likelihood. Even if the two types of structure had an equal prior likelihood in the given context, however, the difference in string edits would still make the posterior probability of the pseudosluice-style structure higher, making it the 'better' choice. Given this large difference in string edits, a P-pied-piping with case-matching continuation would only have a higher posterior probability than a pseudosluice source if cleft-like embedded questions occur much less frequently in such contexts compared to *wh*-movement ones with P-pied-piping.

(105) I Maria chorepse me kapjon, alla de thimame...
*Det.*F.NOM *Maria.*F.NOM *danced.*3SG *with someone.*M.ACC *but* NEG *remember.*1SG..
'Maria danced with someone, but I don't remember...'

a. … pjo**n**.
… *who*.M.ACC
'… who.'

b. … me pjo**n**.
… *with who*.M.ACC
'… with who.'

c. … pjo**n** chorepse me.
… *who*.M.ACC *danced*.3SG *with*
'… who(m) she danced with.'

d. … me pjo**n** chorepse.
… *with who*.M.ACC *danced*.3SG
'… with who(m) she danced.'

e. … [~~me~~] [~~pjon~~] pjo**s**.
… *with who*.M.ACC *who*.M.NOM
'… who.'

f. … me [~~pjon~~] pjo**s**.
… *with who*.M.ACC *who*.M.NOM
'… with who.'

g. … pjo**s** chorepse me.
… *who*.M.ACC *danced*.3SG *with*
'… who(m) she danced with.'

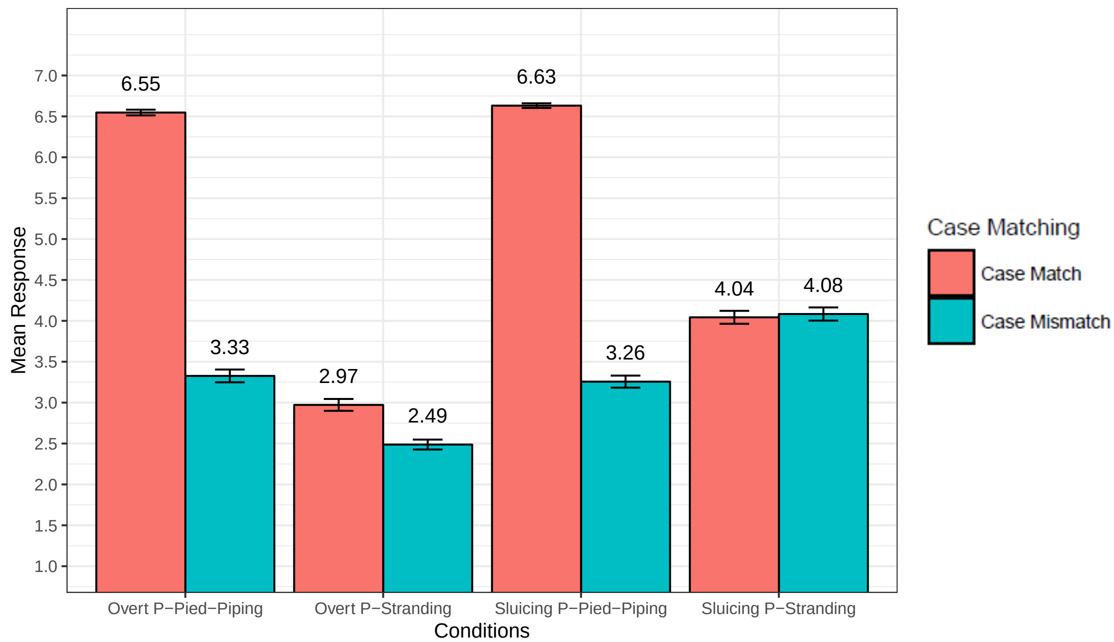
h. … me pjo**s** chorepse.
… *with who*.M.ACC *danced*.3SG
'… with who(m) she danced.'

i. … pjo**s** itan (aftos) me ton opion chorepse.
… *who*.M.NOM *was*.3SG *he*.NOM *with Det*.M.ACC *who*.M.ACC *danced*.3SG
'… who it was with whom she danced.'

The difference between Greek and German may, therefore, be linked to such a difference in structural frequency, i.e. in the prior probability of these structures. If, for instance, cleft-like embedded questions in such contexts are more frequent in German, this may make the German comprehender more likely to opt for such a continuation. In this case, given no further string edits are required, the acceptability of this condition should be equal to that of such an overt pseudosluice-style continuation (once again, assuming no other processing pressures present). If this is true, then the equal ratings of P-stranding with and without case-matching in German would be purely coincidental: the case-matching one reflecting the cost associated with the one string edit required to reach the more acceptable P-pied-piping

Figure 3.6: Experiment 4: Acceptability Ratings Barplot

Mean response per condition with SEM error bars.



with case-matching; the case-mismatching one reflecting the acceptability of a pseudosluice-style continuation. In Greek, on the other hand, such cleft-like embedded questions may be much less frequent when compared to regular wh-movement ones with P-pied-piping. If this is the case, then it is possible that a Greek comprehender is opting for a P-pied-piping case-matching continuation as the most likely intended message for both P-stranding sluicing conditions, with or without case-matching. If this is true, then the difference in Greek acceptability ratings between P-stranding with case-matching vs. with case-mismatching would reflect the different number of string edits required to reach this intended acceptable message.

When considering such frequency-related predictions, as have also been made in previous sections, one way of testing them more thoroughly would be through a corpus analysis, which, however, the time constraints of my PhD did not allow me to undertake. Another way of testing them would be to conduct a sentence completion task involving P-stranding and P-pied-piping remnants with case-matching vs. case-mismatching and observing what types of continuations participants provide and with what frequency. We will consider a form of such a sentence completion task in the next chapter.

3.3.3 Island amelioration under Sluicing

With respect to island amelioration under sluicing, the predictions of a noisy channel model are much more straightforward. In the situations where islands are ostensibly ameliorated, I shall assume here that what is occurring under sluicing is not really island 'amelioration' per se, but rather that apparently island-violating sluices have alternative, non-island-violating sources (Abels, 2011; Abels & Thoms, 2014; M. Barros et al., 2014). This concept was introduced in Chapter 1, with the reader being referred to a more detailed exploration of the topic in (M. Barros et al., 2014). To reiterate here, islands are separated into those which are considered to be ameliorated under sluicing vs. those which are not, such as islands in contrastive sluicing (Fukaya, 2007; Griffiths & Lipták, 2014; Merchant, 2008). It should be noted that this distinction is made in the theoretical literature based on anecdotal evidence, as mentioned previously, with no existing experimental data to date targeting island amelioration in any language beyond those presented here in Chapter 2's Experiment 2 (2.2) for Greek. For this reason, for the time being I will focus on the islands which are argued to be ameliorated, given I have data to support this assumption, leaving those which are supposedly not for further future exploration, once supporting evidence has been collected. For these ameliorated islands, it has been argued that their acceptability derives from non-island-extracting sources. One of these sources is resumption, where a resumptive pronoun is assumed to be contained within the sluice's *e*-site, thereby avoiding island extraction (Abels & Thoms, 2014; Boeckx, 2008; Wang, 2007), as in (106) (M. Barros et al., 2014, (109)). This is the same logic as used to explain apparent P-stranding under sluicing in some overtly non-P-stranding languages, such as Saudi Arabic, as seen in Chapters 1 and 2. The other alternative source is simply a shorter one which contains no island to begin with, as in (107) (Merchant (2001), in part; Abels (2011); M. Barros et al. (2014)). This is also the same logic as used when explaining apparent P-stranding under sluicing in overtly non-P-stranding languages via a cleft or copular source (Rodrigues et al., 2009).

(106) **Resumption (Libyan Arabic)**

miš ʕarəf ʔayya luǵa tanya yəbb-u ywaḋḋf-u šaxṣ
NEG *know*.1.M.S *which* *language* *other* *want*.3.M.PL *employ*3.M.PL. *someone*
yətkalləm=ha
speak.3.M.S-*it*

'I don't know what other language is such that they want to hire someone who speaks it.'

(107) **Short Source**
They wanted to hire someone who speaks a specific language, but I don't remember

which language he should speak.

Given the logic of calling upon such alternative sources to explain island amelioration is the same as that used to circumnavigate potential P-stranding under sluicing in some apparently PSG-defying languages, the predictions we would make here based on a noisy channel approach will be identical to those made earlier for such PSG-defying languages in section (3.3.1.1) of this chapter. To reiterate, a noisy channel would predict that whenever the parser is forced to make a decision amongst possible intended messages, it should always automatically select the one with the highest posterior probability. In this situation, when faced with the structure in (108a), the parser can either assume the completely acceptable structure in (108c) as the intended message, requiring no further string edits; or it can assume the unacceptable structure in (108b) was instead intended, with no string edits required. Given the difference in prior probability between the two options, the posterior probability of the first option is much higher. As such, the parser should theoretically never even consider the second option. Furthermore, given no string edits are required for this alternative source to become acceptable, it should be easily inferred without any additional processing costs. As such, the acceptability rating of this 'island-ameliorating' sluice should be the same as the rating of this alternative source when it is overt (once again, assuming no other processing pressures apply). Furthermore, depending on the rating of the alternative source overtly, the sluicing structure could be as acceptable as other sluices with sources containing no violations.

(108) a. They wanted to hire someone who speaks a specific language, but I don't remember which language.

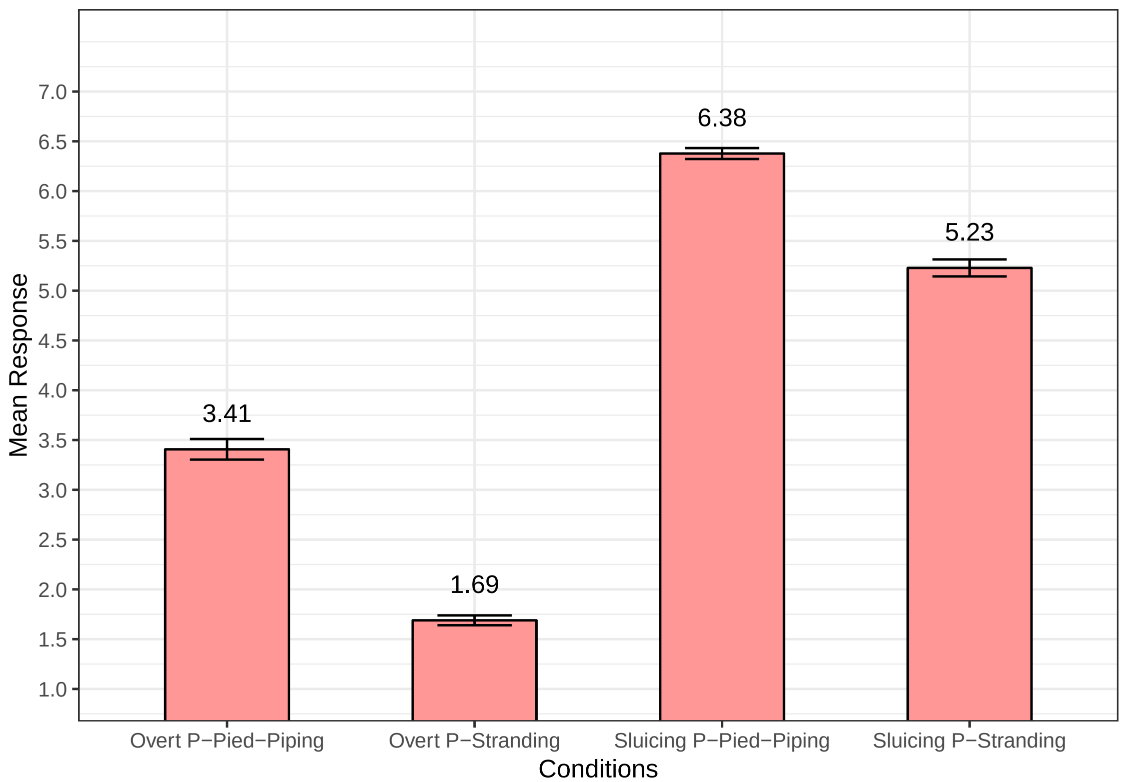
b. * They wanted to hire someone who speaks a specific language, but I don't remember which language they wanted to hire someone who speaks.

c. They wanted to hire someone who speaks a specific language, but I don't remember which language he should speak.

Let us now consider how these predictions fare with respect to the island data from Experiment 2 in Chapter 2 (presented in Figure 3.7). To recap, this experiment crossed the factors of P-stranding and Sluicing in the context of islands (of the subject and relative clause kind), keeping the items as close as possible to those used in Experiment 1, i.e. in the context of no islands, with the same fillers used for both. The results showed that overt island extraction was unacceptable with both P-pied-piping and P-stranding, with the two violations in the overt P-stranding condition appearing to have an additive effect on acceptability[10]. These two overt conditions were also significantly less acceptable than their case-matching counterparts in Experiment 1. Under sluicing, however, both P-stranding and P-pied-piping 'out of an island' became significantly more acceptable, although P-stranding was still significantly less acceptable than P-pied-piping. Comparing these two sluicing conditions with the case-matching sluicing conditions from Experiment 1 (shown again in Figure 3.8), we can see that the two sets of results are very nearly identical. Indeed, an LME model comparing the results of the two experiments, as reported in section 2.2.4, found no significant difference between these sluicing conditions in the two experiments. In other words, P-pied-piping and P-stranding 'out of an island' under sluicing had exactly the same results as if no island were present.

Figure 3.7: Experiment 2: Mean Acceptability Ratings Barplot

Mean response per condition with SEM error bars.

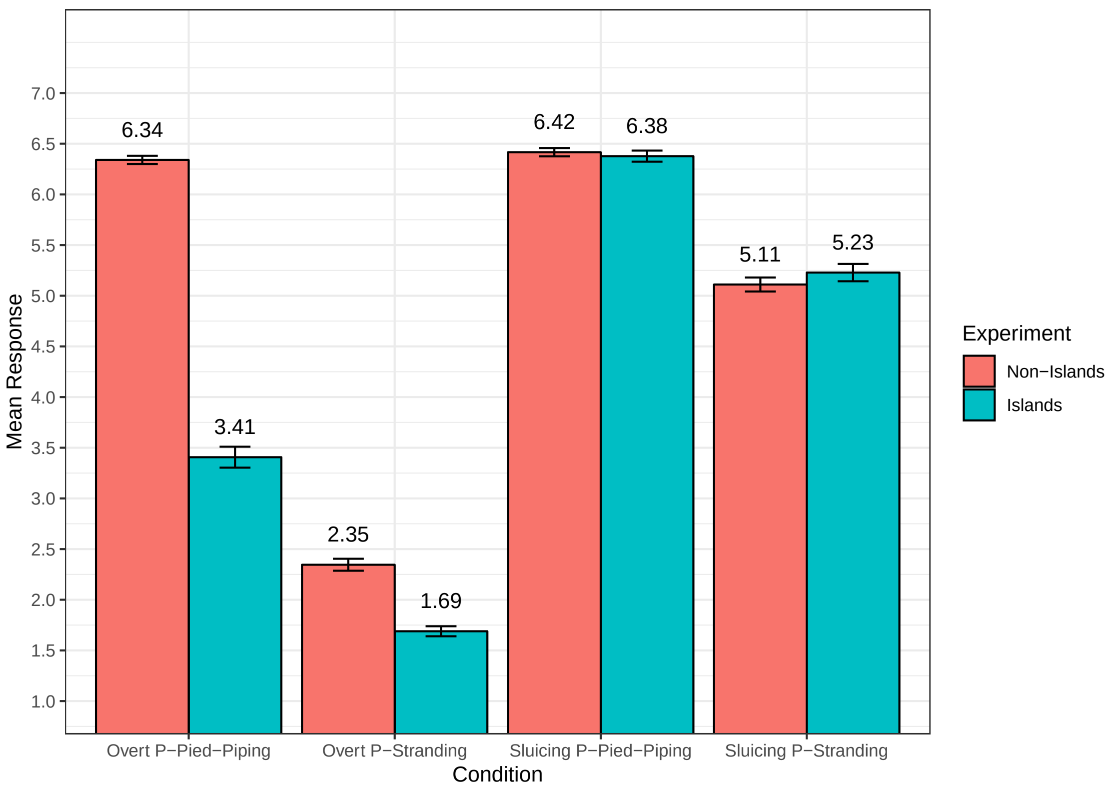


The sluicing results from Experiment 2 and their comparison with those of Experiment 1 make perfect sense based on a noisy channel approach, i.e. it would appear as though

[10]This is very similar to how the two overt violations of case-mismatching and P-stranding had a larger combined effect on acceptability ratings compared to the effects of each separately in the results of Experiment 1.

Figure 3.8: Experiments 1 and 2: Mean Acceptability Ratings Barplot

Mean response per condition with SEM error bars. Only case-matching conditions shown from Experiment 1



the parser is not considering an island-containing source at all, automatically inferring the sluice in (109) as having (109b) as its source instead of (109a). Even the difference between P-pied-piping and P-stranding conditions under sluicing is identical between the two experiments. The explanation for this difference here is also identical to the one detailed earlier on for P-stranding under sluicing outside of the context of islands. That is to say, what appears as P-stranding here, as in Experiments 1, 3 and 4, is in fact a case of P-deletion, with the drop in acceptability between the two conditions being caused by the additional processing cost associated with inferring the intended message being a P-pied-piped structure, i.e. a structure one string edit away from the perceived one.

(109) To oti i dikijoros erjazotan jia kapjous
*Det.*N.NOM *that Det.*F.NOM *lawyer.*F.NOM *was.working.*3SG *for someone.*M.ACC.PL.
sti mafia itan veveos gnosto, alla kanis den
*in.Det.*F.ACC *mob.*F.ACC *was.*3SG *of.course known.*N.NOM *but no.one.*M.NOM NEG
tolmouse na ksestomisi (jia) pjous akrivos ...
*dared.*3SG SUBJ *say.out.loud for who.*M.ACC.PL. *exactly*
'That the lawyer was working for some people in the mob was of course well-known, but no one would dare say who exactly.'

a. ... itan gnosto oti i dikijoros erjazotan (jia).
*... was.*3SG *known.*N.NOM *that Det.*F.NOM *lawyer.*F.NOM *was.working.*3SG *for*
'... it was well-known that the lawyer was working for.'

b. ... erjazotan (jia).
*... was.working.*3SG *for*
'... she was working for.'

3.4 Conclusion

In this chapter, I have outlined the background and theoretical underpinnings of a noisy channel hypothesis of sluicing. Given its lack of overt content and its reliance on overt material as well as pre-existing knowledge and expectations, this approach considers sluicing to be an example of a natural noisy channel. Through this channel, the parser must calculate the most likely intended message on behalf of the transmitter. This calculation is done via Bayesian estimation, with the parser taking into consideration the prior probability of intended messages together with the likelihood that these messages may have been corrupted into the perceived message, with certain corruptions or string edits, such as deletions, being more likely than others, such as insertions, based on the Bayesian size principle. Based on this same principle, fewer string edits are also considered more likely than more string edits.

The combination of prior probability and likelihood of corruption together generate the posterior probability of each message. The message with the highest posterior probability is then assumed to be the intended message.

This sentence processing approach is taken to work in tandem with a theoretical syntax one in order to explain the intriguing behavioural facets of sluicing that a purely sentence processing account or one based on theoretical syntax alone have thus far been unable to capture. Specifically, it is assumed that there is some form of structure within the *e*-site, which exists in a relationship of mutual semantic entailment with the sluice's antecedent. Sluicing remnants that conform to the demands of such structure, had it been overtly expressed, make parsing these sluices straightforward; this makes case-matching remnants easier to parse than case-mismatching ones in case-marking languages; and P-pied-piping remnants easier to parse than P-stranding ones in languages which do not overtly allow P-stranding. That having been said, thanks to the parser's innate ability to calculate the most likely intended message, what have previously been considered 'unacceptable' instances of sluicing based on the above logic, such as instances with case-mismatching or P-stranding remnants, are instead deemed parsable albeit more costly in terms of processing. Specifically, assuming the parser is able to infer a likely intended message following a number of string edits to the perceived one, it will do so, with more string edits being harder to parse. This difference in processing costs translates to off-line decreases in acceptability ratings.

Based on such a combined approach of syntax and sentence processing, a series of cross-linguistic predictions were made with respect the three main characteristics of sluicing, namely the two form-identity generalisations and island amelioration. It should be made clear that each of these sets of predictions is simplified in that it is based solely on the effects that a noisy channel approach alone would predict, i.e. without taking into account other processing effects which have been previously suggested in the sentence processing literature, such as e.g. the repeated name penalty. First and foremost, for P-stranding, we generated a set of cross-linguistic predictions, hypothesizing that in an overtly P-stranding language, both P-pied-piping and P-stranding under sluicing should be as acceptable as overtly - and likely as acceptable as each other. With respect to overtly non-P-stranding languages, a distinction was made; if the language does not have some alternative, non-P-stranding source available, then 'P-stranding' remnants should be treated as cases of P-pied-piping with a deleted P. As such, the string edit required (insertion of a P) in order to reach the intended P-pied-piping message should make these 'P-stranded' remnants more costly to parse compared to P-pied-piping ones, thus making them less acceptable. They should still be significantly more acceptable than overt P-stranding structures, however, thanks to the contextual noise that sluicing introduces compared to no such noise in fully overtly expressed structures. If, on the other hand, the language *does* have an alternative non-P-stranding source available, with no further string edits required, then this source will automatically be inferred with no other calculations being necessary. Given the lack of processing costs associated with this alternative structure, these apparently P-stranded remnants should be as easy to parse as their P-pied-piping counterparts, with no acceptability difference predicted between the two. If, somehow, this alternative source is blocked or is incredibly rare, then as with the other language category above, the parser will resort to treating the 'P-stranded' remnant as a P-pied-piping remnant with a missing P. Inferring the missing P will have an associated cost, however, making P-stranding in this situation less acceptable than P-pied-piping, similarly to the above language category. This means that even within the same overtly non-P-stranding language it is possible to have two different types of 'P-stranding' under sluicing, each with a significantly different predicted acceptability rating, depending on the availability or not of an alternative, non-P-stranding source. These predictions were all confirmed via the experimental data collected here for Greek and German, but also the experimental data collected by AlShaalan and Abels (2019) for Saudi Arabic.

With respect to the second form-identity generalisation, another set of predictions was made. When comparing case-mismatching to case-matching, the former is most obviously a case of replacement of the latter, however a replacement or substitution can also be expressed as a deletion followed by insertion, or vice versa. Treating case-mismatching as an instance of deletion followed by insertion, or vice versa, makes predictions based on a noisy channel approach slightly simpler to visualise. Specifically, if a case-mismatching remnant is parsed as being two string edits away from an acceptable message containing a case-matching remnant, then this would make it harder to parse, and therefore less acceptable, than a case-matching remnant. It should also be harder to parse than P-stranding case-matching remnants, given the latter are only 1 string edit away from an acceptable message. Due to these two string edits, it is unclear whether case-mismatching under sluicing would be easier than case-mismatching overtly.

The only existing experimental data involving case-mismatching under sluicing are those collected here for Greek and German, in Experiments 1 and 4, respectively. These datasets also inevitably included manipulations of P-stranding at the same time as case-matching, as P-stranding behaviour under sluicing was their main target, however they still provide valuable information with respect to case-matching under sluicing. Specifically, results from both languages confirmed that case-mismatching under sluicing was less acceptable than case-matching. They also provided evidence that case-mismatching was more acceptable under sluicing than overtly. Furthermore, in both languages, parsing a case-mismatching remnant with a pied-piped P resulted in *lower* acceptability ratings than one involving P-stranding, even though the latter is theoretically more string edits away from an acceptable P-pied-piping case-matching intended message than the former. This result is considered to be due to the overt P serving as a case-marking reminder directly next to the incorrectly case-marked remnant in P-pied-piping conditions. Results diverged between the two languages, however, with respect to conditions involving both P-stranding and case-mismatching. More specifically, in Greek the combination of the two violations resulted in a lower acceptability rating compared to P-stranding with case-matching under sluicing, as a noisy channel would predict. In German, however, the two conditions were rated as equally acceptable. Although further testing is required to confirm these results, it is possible that there is a different intended message being calculated in each language. In Greek, it is possibly a case-matching P-pied-piping structure which is being inferred, with each string edit bringing an associated processing cost, resulting in the observed difference between P-stranding with and without case-matching. In German, it is possible a different intended message, one compatible with the observed case-marking (Nominative), is being inferred instead, such as a pseudosluice. Given such an intended message would require fewer string edits to be reached from the observed message compared to the one in Greek, it may result in a higher acceptability rating. Its precise acceptability rating would be tied to the number of string edits required and its prior probability of occurring. If this is true, the difference in inferred message between the two languages would likely be tied to the different prior probability of occurence for each structure in each language. That is to say, in German, pseudosluicing may be more easily inferred if cleft-like embedded questions targetting PP-indefinites are more regularly occurring compared to *wh*-questions involving movement, whereas in Greek the opposite may be true.

Finally, with respect to island amelioration, given alternative, non-island-containing sources have time and again been argued to be hiding behind what ostensibly looks like island amelioration under sluicing, adopting such an approach makes predictions based on a noisy channel model simple. Specifically, if the parser can infer an alternative, acceptable source, without any further string edits being necessary, then it will easily do so without incurring any further processing costs. This means that we would predict no acceptability decline between sluices in the context of islands and those in the context of no islands. Indeed, comparing the results of Experiment 1, where there were no islands, to those of Experiment 2, where each antecedent contained an island, confirmed these predictions, finding no significant difference in sluicing results between the two studies.

Having presented this model and its predictions, as well as collected data on the off-line effects of P-stranding on acceptability ratings, the next chapter shall attempt to uncover the on-line processing costs associated with P-stranding in real time. Chapter 4 details these initial attempts to examine the processing cost of P-stranding as compared to P-pied-piping under sluicing through two self-paced reading and one sentence completion study, based on existing literature on the processing of sluicing. Unfortunately, it was not possible to uncover these associated processing costs, as the studies previously proposed in the literature to target sluicing did not prove to be as accurate as argued. Furthermore, examining the way the parser treats PPs showed this behaviour to be more complicated than previously thought, necessitating an alternative approach moving forward.

Chapter 4

Processing Sluicing

4.1 Introduction

In the previous chapters we have examined the theoretical and processing literature concerning P-stranding under sluicing, identifying a gap in the existing literature in terms of experimental data on this phenomenon. We thus conducted a series of off-line acceptability judgement studies identifying precisely how Greek and German speakers rate P-stranding under sluicing compared to overtly and to P-pied-piping. The resulting datasets, however, could not be completely accounted for by any of the existing approaches to sluicing, theoretical or otherwise. In order to successfully capture these data, as well as create a more cohesive cross-linguistic set of predictions, therefore, I proposed a different account, one combining syntax with sentence processing in the form of a noisy channel account of sluicing. The main claim of this account is that the higher acceptability ratings we find associated with P-stranding and/or case-mismatching under sluicing compared to overtly are due to sluicing introducing the contextual possibility of 'noise' for the comprehender. It thus becomes more plausible that the perceived structure may not perfectly match the intended one due to this noise, but may, instead, necessitate some 'tweaking' to arrive at the correct intended structure. This 'tweaking' takes the form of string edits to the perceived structure, with more string edits being harder to parse than fewer string edits. This additional parsing cost is then translated to varying acceptability ratings. The more string edits required for a perceived message to reach an acceptable intended message, the higher the processing cost associated with parsing it and, by extension, the lower its acceptability rating. As such, it is evident that this account of sluicing along with its varying cross-linguistic acceptability rating predictions rely on P-stranding and case-mismatching being costly to parse. In other words, this account predicts an interwoven set of off-line and on-line behaviours, as others have also done for various syntactic phenomena (Hofmeister, 2007; Hofmeister et al., 2014, 2013; Sprouse, 2008; Sprouse et al., 2012).

Thus far, all the previous chapters here have dealt only with the off-line side of P-stranding. In order to see the whole picture of this phenomenon, however, we must also observe it on-line and see whether, indeed, P-stranding under sluicing does carry a larger associated processing cost compared to P-pied-piping in overtly non-P-stranding languages, as our model predicts. Logically speaking, we would anticipate that this additional parsing cost should be evident at the *wh*-phrase or immediately thereafter, with P-less *wh*-phrases being harder to integrate than P-pied-piping *wh*-phrases. This is precisely what the studies in this chapter aimed to do.

Whilst surveying the existing on-line studies of sluicing, a previously conducted self-paced reading study involving sluicing by Yoshida, Dickey, and Sturt (2013) stood out as the ideal experimental set-up to target the potential cost of P-stranding vs. P-pied-piping in Greek. The original study's aim was to examine whether ellipsis, in the form of sluicing, is the parser's preferred economy mechanism and as such is predicted by default wherever possible. Importantly, the study's design made manipulating P-stranding vs. P-pied-piping and even case-mismatching of sluicing remnants relatively simple, as will be seen shortly. However, given the original study was conducted in English, the first step to manipulating the experimental design for our purposes in Greek was to first establish that the original study findings could indeed be replicated in Greek. Upon trying to replicate the study, however, I found some very interesting, different results, suggesting an alternative explanation may in fact underlie the original study findings and that a confounding factor may have been present, one which the nature of Greek made more obvious. In order to further elucidate these new results, I therefore designed and conducted a simplified English version of the Yoshida et al. (2013) experiment, eliminating potential confounds. The results of this simplified self-paced reading study provided further support for the hypothesis that the original study's findings were driven by alternative phenomena. A final English sentence completion task further confirmed this.

As a result of these three studies, it became clear, firstly, that the Yoshida et al. (2013) design was in the end unsuitable for our intents and purposes, but also, secondly, that the generalisations based on those original study results, and other similar designs since, appear to be incorrect and should, therefore, be reconsidered. Although the original goal of the studies presented in this chapter was not reached, i.e. to target the various processing costs hypothesized to be associated with P-stranding and case-mismatching, we at least determined ways in which we should *not* move forward with this investigation. One important finding, in particular, that needs to be taken into consideration moving forward is that the way in which we process PPs is, in fact, much more complicated than the literature has thus far considered. Specifically, the parser appears to only sometimes take into consideration the presence or absence of these Ps in predicting likely intended messages, whilst at other times it completely ignores them. As such, using Ps to guide the parser towards or away from certain structural predictions does not seem to work as well as we might expect. Whilst generally very interesting, this finding also adds further evidence to the idea that our previous expectations of upcoming material can shape how we perceive bottom-up input - and the presence or absence of Ps in particular - in order to generate the most likely and cognitive effort-conserving intended message. Although this is not the direct evidence we were looking for in favour of a noisy channel hypothesis, this finding does seem to be indirectly consistent with it.

4.2 Background: Yoshida, Dickey & Sturt, 2013

4.2.1 Study Design & Results

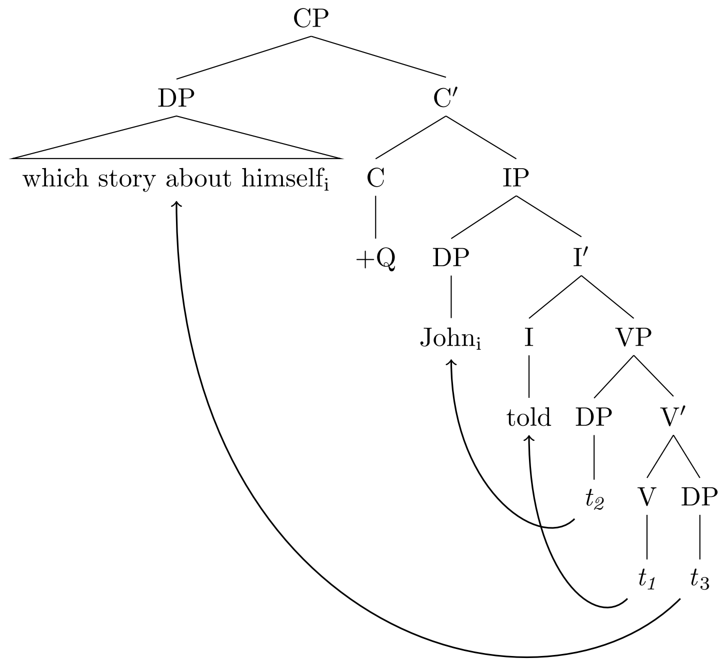
In 2013, Yoshida et al. published a very interesting study on sluicing prediction. The aim of this study was twofold: to show that the *e*-site within the sluice contained fully detailed hierarchical syntactic structure; but also that we have an innate tendency to predict sluicing - and by extension ellipsis - whenever the opportunity presents itself.

The foundations of the study lie in the existing processing literature claiming that the parser actively generates predictions about upcoming stuctures based on a combination of already available syntactic material along with grammatical knowledge and restrictions (E. Lau, Stroud, Plesch, & Phillips, 2006; Phillips, 2006; Staub & Clifton Jr, 2006; Wagers & Phillips, 2009). Specifically, the experimental manipulation is based on two pre-existing effects, one - arguably - syntactic in nature and one from the sentence processing literature. The first is the so-called *anaphor connectivity effect* (Merchant, 2005b; Stjepanović, 2008), i.e. the fact that sluicing appears to allow binding of a reflexive by an antecedent external to its binding domain (110a), contrary to Binding Condition A (Chomsky, 1981) (110b).

(110) a. John$_i$ told some stories at the family reunion [$_{\text{CP}}$but we could not remember [$_{\text{CP}}$which story about himself$_i$]].
b. John$_i$ told some stories at the family reunion [$_{\text{CP}}$but we could not remember [$_{\text{CP}}$which story about himself$_{*i/j}$ William$_j$ told]].

This effect has been explained by positing that the *e*-site contains fully-fledged hierarchical syntactic material (ignoring for the moment exactly how that material appears there). In this way a (form of) syntactic copy of the sluice antecedent containing the anaphor antecedent is available within the anaphor's binding domain to grammatically bind it without being explicitly pronounced (e.g. (111), with the relevant *e*-site structure shown, as a source for (110a)). This explanation is obviously compatible with a syntactic analysis of the *e*-site (e.g. Merchant (2001, 2005b); Ross (1969), a.o.).

(111) but we could not remember...



The second effect that the manipulation takes advantage of is a robust processing one found repeatedly for reflexive anaphors and pronouns in self-paced reading and event-related potential (ERP) studies, namely the gender mismatch effect (henceforth 'GMME') (Badecker & Straub, 2002; Kazanina et al., 2007). This effect arises specifically when an anaphor does not immediately detect a licit potential antecedent (restricted by the rules and conditions of binding theory) which matches it in gender, but instead finds one which does not. The result is a measurable slow-down in participants' reaction times compared to when the possible antecedent and anaphor match in gender. If potential antecedents precede the anaphor (anaphora), the effect occurs at (or just after) the anaphor region, whereas if they follow it (cataphora), the effect will occur at (or just after) the potential antecedent.

These two effects are combined to create an elegant study design, albeit with complex stimuli. This complexity will play an important role in a later argument I will present regarding the interpretation of the study's results. Starting with the stimulus in (112a), Yoshida et al. manipulate the factor of reflexive gender matching, by changing the first sentence's subject to either match the reflexive in gender (112a) or mismatch it (112b) (with the two relevant elements underlined).

(112) a. Marie's grandfather told some stories at the family reunion, but we couldn't remember which story about himself ...

b. Marc's grandmother told some stories at the family reunion, but we couldn't remember which story about himself ...

The key here is that both (112a) and (112b) are compatible with a sluice up until the reflexive region, but only (112a) remains compatible once the reflexive is encountered due to the mismatch in gender. If, therefore, a sluice is our preferred continuation, meaning that we project it in advance of bottom-up confirmatory material in anticipation, and if, furthermore, this sluice projection contains full hierarchical structure abiding by the rules of binding theory, then we would expect a GMME at the reflexive in (112b) as per the anaphor connectivity effect found in sluicing.

In order to ascertain that such a mismatch effect would indeed be due specifically to a sluice being projected, the additional factor of sluice compatibility is cleverly introduced. Specifically, a pair of stimuli minimally different to those in (112), but which do not allow a sluice projection, are created by adding a P incompatible with the main clause verb directly before the *wh*-phrase, such as *with* here (highlighted in bold in (113)). This P creates the appearance of a full pied-piped PP containing the *wh*-element (i.e. *with which story*), however this pied-piped PP is incompatible with the main clause's verb (**told with which story*). One must bear in mind here that if at this point the parser is projecting a sluice, this *wh*-XP would serve as the sluice's remnant with the main clause as its antecedent. This, however, also entails that the remnant would need to be compatible with its antecedent's verb (114). Since this verb is not compatible with a PP complement, this P effectively removes the option of a sluice projection, making these two additional conditions non-sluice-compatible. If, therefore, a GMME found at the reflexive is due to a sluice being projected and this sluice is unavailable when we have a wh-PP, then we would expect to find a reaction time slow-down in (112b), but not in (113b).

(113) a. Marie's grandfather told some stories at the family reunion, but we couldn't remember **with** which story about himself ...

b. Marc's grandmother told some stories at the family reunion, but we couldn't remember **with** which story about himself ...

(114) * Marie's grandfather told some stories at the family reunion, but we couldn't remember **with** which story about himself ~~[he told]~~.

In order for the introduction of this P to result in an overall grammatical stimulus, the structure in (113) is continued as in (115), i.e. using a verb compatible with this PP (e.g. *impressed with*). In order for all stimuli to be grammatical, an antecedent matching the reflexive in gender is also introduced cataphorically within the reflexive's binding domain (*his brother*). To keep all stimuli minimally different and containing the same lexical items, the sluice-compatible examples are also continued in the same way, with the final conditions shown in (116). To make condition names less verbose, the original study dubbed sluice compatibility as 'Wh-Type', with sluice compatible conditions named 'Wh-NP' conditions and sluice non-compatible ones named 'Wh-PP' based on the (potential) remnant's form. As is evident from the set of conditions in (116), no experimental condition actually contained a sluice.

(115) Marie's grandmother told some stories at the family reunion, but we couldn't remember **with** which story about himself from the party his brother was so very impressed.

(116) a. **Condition 1. (Gender Matching; Wh-NP)**
Marie's grandfather told some stories at the family reunion, but we couldn't remember which story about himself from the party his brother was so very impressed **with**.

b. **Condition 2. (Gender Mismatching; Wh-NP)**
Marc's grandmother told some stories at the family reunion, but we couldn't remember which story about himself from the party his brother was so very impressed **with**.

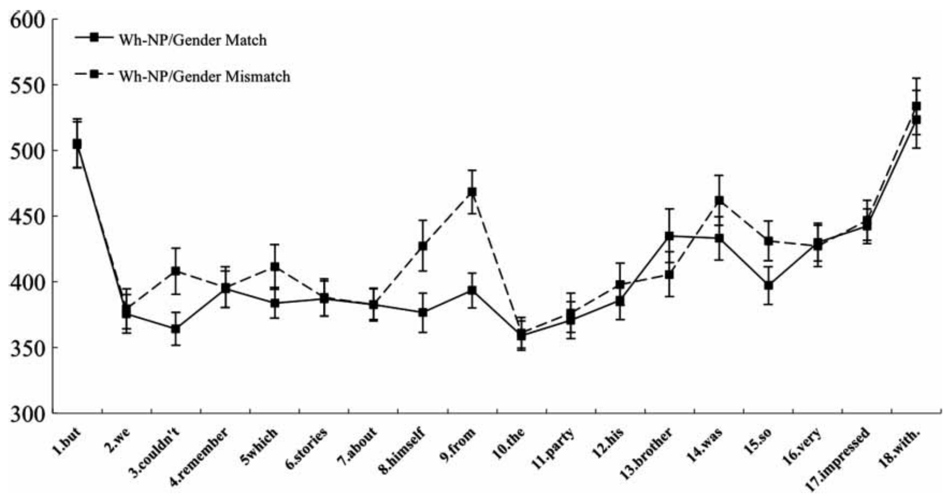
c. **Condition 3. (Gender Matching; Wh-PP)**
Marie's grandfather told some stories at the family reunion, but we couldn't remember **with** which story about himself from the party his brother was so very impressed.

d. **Condition 4. (Gender Mismatching; Wh-PP)**
Marc's grandmother told some stories at the family reunion, but we couldn't remember **with** which story about himself from the party his brother was so very impressed.

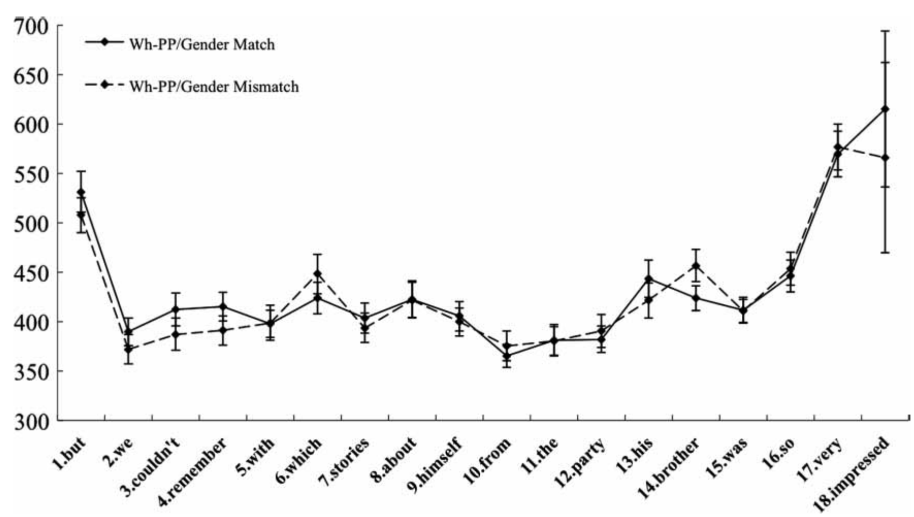
Based on the experimental design above, Yoshida et al. (2013) have a clear experimental hypothesis regarding sluice prediction: if, whenever possible, we project fully hierarchical sluices in advance of bottom-up confirmation, we would expect to find a GMME at the reflexive (*himself*) in the gender mismatching, sluice compatible condition only, and not in the sluice incompatible one. In other words, we would expect a significant interaction between the two factors of Gender Matching (±) and Wh-Type (Wh-NP vs Wh-PP) with Gender Mismatching creating a significant slow-down in reading times at the reflexive region (or just beyond it, due to spill-over effects) *only* when we have a Wh-NP (Condition 2), and not when we have a Wh-PP (Condition 4), thanks to the latter blocking a sluice continuation. If, on the other hand, this sluice projection does not occur or is not as detailed, then we would expect no such interaction nor any main effects.

This predicted interaction is exactly what Yoshida et al. (2013) found, with the SPR study's mean reading time results shown in Figures 4.1a and 4.1b (their figures (1) and (2) respectively). The region of interest here is the reflexive '*himself*' and the immediately following region, '*from*', (regions 8 & 9 in (4.1a) and 9 & 10 in (4.1b)). All regions were analysed using a repeated measures ANOVA, with reflexive Gender Matching (Matching vs Mismatching) and Wh-Type (NP vs. PP) as within-subjects and within-items factors.

Figure 4.1: All Yoshida et al. (2013) SPR results



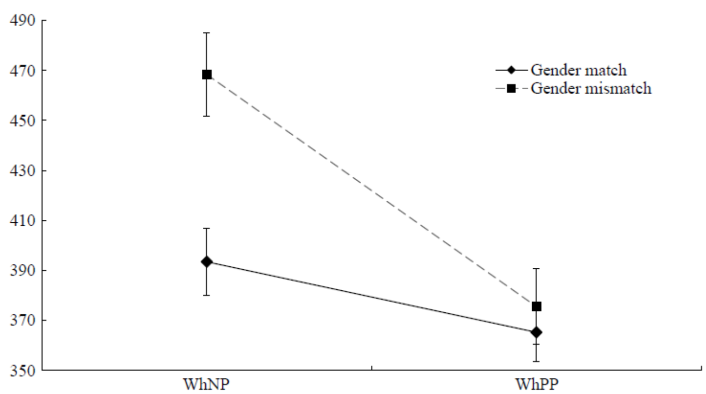
(a) WhNP Results



(b) WhPP Results

The ANOVA showed a significant main effect of Gender Matching at the reflexive itself with gender mismatching in the Wh-NP condition being read significantly more slowly than in the other conditions (significant only by participants (F1) and not items (F2); $F1(1,31) = 5.00, p < .05; F2 < 1$), although apparently the region showed no significant interaction[1]. Pairwise comparisons[2] at the reflexive found a significant difference between the two Wh-NP conditions (gender mismatch significantly slower than gender matching) ($F1(1,31) = 5.68, p < .05; F2(1,23) = 4.03, p = .057$), with no such difference between the two Wh-PP conditions ($F's < 1$). For the region immediately following the reflexive, i.e. the preposition '*from*', results showed main effects of both Gender Matching ($F1(1,31) = 10.35, p < .01; F2(1,23) = 6.47, p < .05$) and Wh-Type ($F1(1,31) = 10.75, p < .01; F2(1,23) = 5.22, p < .05$), with these effects being modulated by a significant interaction between factors ($F1(1,31) = 11.33, p < .01; F2(1,23) = 6.47, p < .05$). Figure (4.2) (their Fig. 3) illustrates the interaction in this region. Pairwise comparisons showed that for the Wh-NP conditions, the gender mismatching condition was read significantly more slowly than the matching one ($F1(1,31) = 31.71, p < .01; F2(1,23) = 15.37, p < .01$), whereas this was not the case for the Wh-PP conditions ($F1(1,31) = 0.01, p = .9; F2(1,23) = 0.7, p = .41$).

Figure 4.2: Interaction Summary



Yoshida et al. (2013) take these results as supporting evidence for their experimental hypothesis that participants are indeed able and inclined to project a sluice continuation whenever possible, i.e. in this situation once they reach the *wh*-phrase. Moreover, this projection is detailed enough to cause a GMME when the potential antecedent of a sluice-contained reflexive does not match this reflexive in gender.

These on-line results were reportedly further bolstered by two follow-up off-line studies. Specifically, in order to establish that the Wh-NP conditions were indeed compatible up

[1]NB: This result seems unexpected, given the results report and the graphs.

[2]NB: These were presumably adjusted for multiple comparisons, although this is not clarified.

until the reflexive region with a sluice continuation, but that the Wh-PP conditions were not, an acceptability judgement task and a sentence completion (SC) task were performed. For the acceptability judgement task, truncated versions of these four conditions were used. These truncated versions contained the full condition up until the reflexive, as seen in (117). The idea here is that the Wh-NP conditions should be significantly more acceptable than the Wh-PP ones, as the former are supposed to be compatible with a sluice, whereas the latter are not[3].

(117) a. **Condition 1. (Gender Matching; Wh-NP)**
Marie's grandfather told some stories at the family reunion, but we couldn't remember which story about himself.

b. **Condition 2. (Gender Mismatching; Wh-NP)**
Marc's grandmother told some stories at the family reunion, but we couldn't remember which story about himself.

c. **Condition 3. (Gender Matching; Wh-PP)**
Marie's grandfather told some stories at the family reunion, but we couldn't remember with which story about himself.

d. **Condition 4. (Gender Mismatching; Wh-PP)**
Marc's grandmother told some stories at the family reunion, but we couldn't remember with which story about himself.

Table 4.1: Mean Acceptability Ratings per Condition

	Condition	Rating
1	Gender Matching, Wh-NP	3.35/5.00
2	Gender Mismatching, Wh-NP	2.15/5.00
3	Gender Matching, Wh-PP	1.84/5.00
4	Gender Mismatching, Wh-PP	1.72/5.00

The study's Mean ratings per condition are shown in Table (4.1). With a scale of 1 to 5 (1: completely unacceptable; 5: completely acceptable), results showed that on average, participants rated the Wh-NP conditions significantly more highly than the Wh-PP ones (2.75 vs. 1.99 respectively; $F1(1,39) = 76.07, p < .001; F2(1,23) = 72.39, p < .001$). Moreover, there was a significant interaction, with gender mismatching having a significant

[3]It should be noted here, however, that the Gender Mismatching Wh-NP condition does not actually contain a grammatical antecedent for the reflexive, thereby rendering it unacceptable. As such, it is not made clear why a main effect is expected instead of an interaction.

effect only on the Wh-NP conditions, and not the Wh-PP ones ($F1(1, 39) = 69.58, p < .001; F2(1, 23) = 69.27, p < .001$).

For the SC task, participants were asked to complete truncated versions of the same SPR stimuli, this time truncated just before the reflexive, as in (118), meaning that Gender Matching no longer constituted a factor. The idea behind this task is that if the pied-piped P in the Wh-PP conditions is truly incompatible with a sluice, then we would anticipate significantly more sluice completions for the two Wh-NP conditions compared to the two Wh-PP ones.

(118) a. **Condition 1. (Wh-NP)**
Marie's grandfather told some stories at the family reunion, but we couldn't remember which story about ____.

b. **Condition 2. (Wh-NP)**
Marc's grandmother told some stories at the family reunion, but we couldn't remember which story about ____.

c. **Condition 3. (Wh-PP)**
Marie's grandfather told some stories at the family reunion, but we couldn't remember with which story about ____.

d. **Condition 4. (Wh-PP)**
Marc's grandmother told some stories at the family reunion, but we couldn't remember with which story about ____.

Completions consisting of a 'single noun or noun phrase' (p. 12) were considered *Sluice Completions*, whereas all other completions (blanks not accepted) were marked as *Other*. Out of the 864 completions collected, sluice completions made up 65.1% of Wh-NP condition completions and 15.7% of Wh-PP completions (with mixed logistic regression modelling showing a reliable effect of Wh-Type: $Z = 11.80, p < .001$).

The authors took the results of both of these follow-up studies to indicate that their SPR stimuli did indeed create the distinction between sluice compatibility for the Wh-NP conditions and sluice incompatibility for the Wh-PP conditions. Furthermore, the acceptability data was taken as further evidence that this sluice projection is detailed enough to cause a GMME for a reflexive contained within this projected sluice. Together with the results of the SPR study, they conclude that not only do we have an innate tendency to project sluicing, and by extension ellipsis, whenever possible, but also that these projections are detailed and hierarchically organised, abiding by the rules of binding theory.

4.2.2 Objections

Although the Yoshida et al. (2013) on-line study design is undoubtedly elegant and the results of both this and the two off-line studies do appear to support the authors' experimental hypothesis, there are a few objections which deserve to be raised at this point.

The first objection has to do with the SPR study's design revolving around a reflexive, particularly one contained within what is known as a 'representational nominal' or 'RNP', such as *picture of him/himself.* Starting with the issue of the reflexive itself, one possibility which may have occurred during the task is that encountering the reflexive alone may have been in and of itself a sufficient enough cue to generate a search for the reflexive's antecedent, without necessarily presupposing the projection of a sluice. The main argument which could be levelled against this is that we have evidence that reflexives do indeed create GMME, but only for grammatically licensed antecedents (Kazanina et al., 2007; Sturt, 2003) (though cf. Badecker and Straub (2002); Runner et al. (2003) for evidence of interference from grammatically non-licensed potential antecedents). If we accept this as true, then out of the SPR conditions, only the Wh-NP ones provide a grammatically licensed antecedent through a sluice projection. This brings me to the next important objection, however, and that is that even if this is true, pronouns, including reflexives, contained within these 'RNPs' (e.g. *picture of him/himself, drawing of him/himself* etc.), like those used in this study, have repeatedly been claimed to be exempt from the constraints of binding theory (Kaiser, Runner, Sussman, & Tanenhaus, 2005; Pollard & Sag, 1992; Reinhart & Reuland, 1993; Runner & Goldwater, 2011). Although an explanation of this behaviour is beyond the scope of this discussion, with the reader referred instead to Pollard and Sag (1992); Reinhart and Reuland (1993); Runner and Goldwater (2011); a.m.o., it has been shown several times that these RNP-contained reflexives, as opposed to regular reflexives, are able to break the rules of binding. Firstly, they do not appear to require a c-commanding antecedent (119); secondly, they may take an antecedent outside their minimal binding domain (i.e. the clause) (120); and thirdly, they even appear able to take an antecedent in a completely separate sentence ((121) (Pollard & Sag, 1992)).

(119) a. * Norman$_i$'s sister bought himself$_i$ some new shoes.

b. That picture of himself$_i$ on Facebook had shaken Norman$_i$ quite a bit.

(120) a. Norman$_i$ realised [$_{\text{CP}}$ that John$_j$ bought himself$_{*i/j}$ some new shoes].

b. Norman$_i$ realised [$_{\text{CP}}$ that there was [$_{\text{NP}}$ a picture of himself$_i$] on Facebook].

(121) John$_i$ was going to get even with Mary. That picture of himself$_i$ in the paper would really annoy her, as would the other stunts he had planned.

If this is so, then this renders null and void the study's central concept of hierarchical sluice projections being the only available reflexive licensors in these conditions. In other words, finding a GMME at the reflexive need not necessarily indicate the projection of fully hierarchical structure at all, but simply that the reflexive itself is generating a search for its antecedent, independently of the surrounding context.

In their Discussion section, Yoshida et al. do indeed acknowledge this potential issue; their response is that if this were indeed the driving force behind the mismatch effect, then we would not expect to find an interaction between the availability or not of a sluice and gender matching at the post-reflexive region (*from*). Instead, if it is the reflexive which is independently generating a search for its antecedent, then we would anticipate a main effect of gender matching, with gender mismatching always incurring a reaction time slow-down, regardless of sluice compatibility.

Although the logic of this statement is clearly sound, there is one important further objection which should be raised at this point and which has been overlooked by the authors. Specifically, in the Wh-PP conditions, the stimulus structure allows for a local attachment ambiguity of the pied-piped P (*with*). This P can be interpreted, for instance, as a) the start of an Adverbial Phrase which may typically modulate the previous verb, e.g. (122), b) the start of the previous verb's complement, e.g. (123), or c) the beginning of a passive voice clause, e.g. (124).

(122) Marc's grandmother told some stories at the family reunion, but we couldn't remember <u>with certainty/clarity</u>...

(123) Marc's grandmother told some stories at the family reunion, but we didn't remember <u>to record them</u>...

(124) Marc's grandmother brought some books to the auction, but we couldn't remember <u>by which authors they were written</u>...

This possibility for misinterpretation becomes all the more likely when we take into consideration that in English embedded clauses, it has been argued that pied-piping of Ps is generally dispreferred in favour of P-stranding (Cable & Harris, 2011; Heck, 2008). If this is true, then this makes it more likely that, upon encountering the pied-piped PP, the parser does not interpret it as the argument of an as-yet-unseen verb, but rather attempts to interpret it in one of the more frequently occurring above ways (122, 123, or 124). When the parser realises this is not the case, it needs to alter its full parse predictions, either by re-ranking or re-weighting parallel predictions in a parallel processing model (as is, for instance, the Bayesian estimation model argued for in this thesis) or selecting a different interpretation in a serial processing model. Regardless of the approach taken, however, this recalibration is thought to be associated with increased cognitive demands, resulting in a slow-down or classic garden-path effect. The resulting confusion may mean that the reflexive gender mismatch goes unnoticed in the Wh-PP condition, with this recalibration process being the true root of the observed interaction - an idea we will revisit later on in this chapter when dissecting the follow-up Greek and English study results reported here. To thoroughly test for this possibility, one would have to preclude this attachment ambiguity and see if the study results change in favour of a gender mismatch main effect. Another way to investigate this possibility would be to check the original study results for order effects, i.e. to examine whether participants changed their behaviour over time; by being exposed to more and more instances of P-pied-piping during the experiment, they may have become more used to incorporating the pied-piped P as part of an upcoming verb instead of one of the above interpretations which they may have originally been more inclined towards. By becoming more used to P-pied-piping, this would lead to the reflexive gender mismatch being more readily processed even in the Wh-PP condition, i.e. we may see a potential increase in GMMEs for the Wh-PP condition in the second half or latter third of the experiment. That having been said, the lack of any true sluices in the experiment design means that it is also likely that participants would similarly become more familiar with the reflexive never referring back to the antecedent, matching or otherwise, which would also result in a general lessening of a GMME across all mismatching conditions, meaning the above order effect may be lost.

Setting these objections regarding the reflexive aside, however, another important issue to consider at this point is the accuracy of the study premise in and of itself, i.e. whether this Wh-PP manipulation truly creates a distinction *only* between a) a sluice projection (Wh-NP conditions) and b) a non-sluice projection (Wh-PP conditions). If one is to claim a preference for sluicing based only on this manipulation, then a sluice continuation should clearly be the only option available with sufficient structure to allow appropriate binding of the reflexive. However, as was clear also from the sentence completion follow-up study, this is obviously not the only possible continuation, with more than a third of Wh-NP completions being something *other* than a sluice. For instance, another highly plausible continuation for the Wh-NP conditions which would also give rise to the same interaction found in the Yoshida et al. study would be a simple parallel structure (125) or a cleft (126).

(125) a. Marc's grandmother told some stories at the family reunion, but we couldn't remember which stories about herself she/his grandmother told.

b. Marie's grandfather told some stories at the family reunion, but we couldn't remember which stories about herself she/his grandmother told.

(126) a. Marc's grandmother told some stories at the family reunion, but we couldn't remember which stories about herself it was that she/his grandmother told.

b. Marie's grandfather told some stories at the family reunion, but we couldn't remember which stories about herself it was that she/his grandmother told.

Furthermore, approximately 15% of the Wh-PP completions were sluices, indicating that participants may actually sometimes be entertaining a sluice completion even for the Wh-PP conditions, again making the distinction between Wh-NP and Wh-PP less clear-cut. Although this is not a high percentage, it is still worth taking into account as it may indicate that participants are either misreading the item as a Wh-NP condition, or, perhaps more intriguingly, they may be (actively) overlooking the P - a concept I will also return to when discussing the follow-up data in section 4.3. In order to claim a sluicing default preference, one must, therefore, address these issues, replicating these findings when all other such continuations have been blocked. Although this is impossible in non-case-marking languages such as English, a richer feature-marking language such as Greek can make this distinction. For instance, case-marking of the 'remnant' would automatically disallow a cleft continuation, whilst P-stranding of this same 'remnant', based on our Greek acceptability judgement findings, would likely encourage only a sluice and not a parallel continuation prediction.

Finally, it is also unclear whether the projected sluices containing the reflexive, as in (127), are actually acceptable, with a difference in number-marking between correlate and remnant in some stimuli, as in (128), appearing to add to this potential problem. My own judgements aside, the Mean rating in the Yoshida et al. (2013) acceptability judgement task for the Gender Matching Wh-NP condition (3.35/5.0) was still much closer to a midrange undecided rating (3.0/5.0) than to a perfectly acceptable one (5.0/5.0). Although this rating could have been due to the types of fillers used (which are not described), it certainly indicates that a sluice continuation for this condition, though better than for the other conditions, is still far from perfect.

(127) Marc's grandmother told some stories at the family reunion, but we couldn't remember which stories about herself.

(128) Nathan's mom saw lots of videos at the office party but no one was sure which video of herself.

4.2.3 Conclusion

In conclusion, the SPR study conducted by Yoshida et al. (2013) crossing the factors of Sluice Compatibility and Reflexive Gender Matching provided some evidence for participants projecting a sluice once they have reached the *wh*-XP in a sluice-compatible stimulus, but not in a sluice-incompatible one. This lead the authors to claim that we have an innate preference for sluicing, and by extension ellipsis, when the context allows for it. Two follow-up offline studies provided further data on a) how acceptable these sluice continuations were for what were considered 'sluice-compatible' and 'sluice non-compatible' conditions; and b) how likely participants were to complete each of these conditions truncated at the *wh*-XP with a sluice vs other continuations. Their acceptability results indicated that, although not perfectly acceptable, sluice compatible conditions were more highly rated with a sluice continuation vs. sluice non-compatible conditions; they also found that the former were more likely to be completed with a sluice in a SC task vs. the latter.

That having been said, a number of issues were also raised with respect to the experimental design, particularly regarding a) the inclusion of an RNP reflexive being pivotal, but also b) the design's reliance on a pied-piped PP to create a minimal difference between sluice compatibility vs. sluice incompatibility in conditions. With respect to the former issue, the design relies on anaphor binding restrictions and the presence of a GMME to indicate the prediction of sluicing. However, it has been previously argued repeatedly that RNP reflexives do not necessarily abide by the normal rules of binding. With respect to the latter issue, the first problem is that it is possible that a local attachment ambiguity inherent in the stimulus design may be obscuring the true results of the reflexive gender mismatch manipulation. Secondly, the English language and the way the stimuli are designed currently makes it impossible to distinguish between possible, plausible projections of a) a sluice, b) a parallel structure, and c) a cleft, something which weakens the study's claims.

4.3 Self-Paced Reading Replication in Modern Greek;

4.3.1 Experimental Outline, Aim and Predictions

As mentioned earlier, the reason the Yoshida et al. (2013) study was examined in such detail is because its design holds great potential to allow us to investigate various stages of sluice processing as reflected by reading times. If we are to argue that off-line acceptability differences between P-stranding vs. P-pied-piping as well as case-mismatching vs. case-matching are each tied to increased processing costs in the former vs. the latter, then measuring these costs on-line and further manipulating them to observe their effects on acceptability is a necessary next step to confirm this hypothesis. The Yoshida et al. (2013) SPR study design would make it possible to compare these to detect, for instance, the potential processing cost associated with integrating a P-stranded remnant under sluicing in Greek vs. a P-pied-piping one. It could also be used to examine any factors which may be argued to affect this processing cost, such as the complexity of remnant and/or correlate, as e.g. a pointer account may predict (even if we did not find any off-line evidence for this factor being important).

In order to successfully manipulate the original study design in Greek, however, to examine these costs, it was first necessary to demonstrate it could be replicated in this language. Furthermore, the nature of Greek with its rich feature-marking system makes it the perfect candidate to address some of the potential issues raised with respect to the original study's design and results. The first advantage it presents is that the problematic local attachment ambiguity found for the pied-piped PP in English is not present in Greek. Instead, this pied-piped PP can only be interpreted as the argument of an as-yet-unseen verb. Comparing the Greek results to those of the original English study could, therefore, shed more light on whether the interaction of factors observed at the reflexive in English may have been partially affected by this attachment ambiguity. The second advantage Greek presents is that by overtly case-marking the *wh*-phrase, it is possible to block at least one alternative projection other than a sluice. Specifically, case-marking of the remnant in Accusative can serve to block potential cleft predictions, given the latter, as previously mentioned, must always appear in Nominative. Although the current design does not also block potential parallel structure projections, this is something that could be achieved in follow-up studies by including the factor of P-stranding.

The aim of the first SPR study presented below was, therefore, to replicate the original, and by eliminating certain confounding factors provide further evidence that we do indeed have a default tendency to predict sluicing.

4.3.2 Methods

4.3.2.1 Experimental Design and Predictions

To replicate the original study design, the same factors were manipulated, namely Reflexive Gender Matching (±; Matching vs. Mismatching) and Sluice Compatibility(±). Given sluice compatibility is manipulated through the presence of a pied-piped PP and to keep factor names consistent with those presented above for the original study, the factor of sluice compatibility is referred to here as 'Wh-Type', with Wh-NP conditions being compatible with a sluice projection and Wh-PP conditions not. These two factors were crossed to create the 2x2 table of conditions shown in Table 4.2 (with superscript numbers depicting sentence regions - see also next section).

The main idea here, as in the original, is that if we project a sluice continuation by default, then encountering a reflexive whose gender is dissonant with this projection should create a processing slow-down, i.e. we are interested in observing the presence or not of a significant difference between gender-matching vs. gender-mismatching conditions at or just after this sentence point. In order to establish that this difference is indeed due to a potential sluice projection, however, and not some other driving factor, the presence or not of the pied-piped PP and its effect on this difference is key. Specifically, given there is no local attachment ambiguity of the pied-piped PP, encountering it here should serve to block or foil a sluice or parallel structure projection. As such, if we indeed create such projections in

Table 4.2: Greek SPR Conditions and Examples

	Condition	Example
1	Gender Matching, Wh-NP	The.F.NOM sister.F.NOM the.M.GEN Charis.M.GEN presented many.M.ACC.PL. paintings.M.ACC.PL. at.the exhibition,[1] but[2] NEG[3] was[4] clear[5] which.M.ACC[6] paintings.M.ACC.PL.[7] the.M.GEN[8] self.M.GEN[9] hers[10] from[11] the[12] exhibition[13] was.pleased[14] the.N.NOM[15] model.N.NOM[16] quietly.[17]
2	Gender Mismatching, Wh-NP	The.M.NOM brother.M.NOM the.F.GEN Chara.F.GEN presented many.M.ACC.PL. paintings.M.ACC.PL. at.the exhibition,[1] but[2] NEG[3] was[4] clear[5] which.M.ACC[6] paintings.M.ACC.PL[7] the.M.GEN[8] self.M.GEN[9] hers[10] from[11] the[12] exhibition[13] was.pleased[14] the.N.NOM[15] model.N.NOM[16] quietly.[17]
3	Gender Matching, Wh-PP	The.F.NOM sister.F.NOM the.M.GEN Charis.M.GEN presented many.M.ACC.PL. paintings.M.ACC.PL. at.the exhibition,[1] but[2] NEG[3] was[4] clear[5] for[6] which.M.ACC[7] paintings.M.ACC.PL[8] the.M.GEN[9] self.M.GEN[10] hers[11] from[12] the[13] exhibition[14] was.pleased[15] the.N.NOM[16] model.N.NOM[17] quietly.[18]
4	Gender Mismatching, Wh-PP	The.M.NOM brother.M.NOM the.F.GEN Chara.F.GEN presented many.M.ACC.PL. paintings.M.ACC.PL. at.the exhibition,[1] but[2] NEG[3] was[4] clear[5] for[6] which.M.ACC[7] paintings.M.ACC.PL[8] the.M.GEN[9] self.M.GEN[10] hers[11] from[12] the[13] exhibition[14] was.pleased[15] the.N.NOM[16] model.N.NOM[17] quietly.[18]

Translation: 'Charis's sister / Chara's brother presented many paintings at the exhibition, but it was not clear which paintings of herself the model was secretly pleased with.'

Explanatory Notes: 1. For the examples, a simplified gloss version with only the relevant case-marking is given. 2. Superscript numbers denote presented regions (with the antecedent being presented all at once).

advance of bottom-up material and use cues such as PPs to block potential parses, we would expect to find a significant interaction of Wh-Type with Gender Matching at the reflexive (or the immediately following regions due to spill-over effects), with a larger difference being observed between gender-matching and mismatching Wh-NP conditions compared to Wh-PP conditions, where such projections are blocked. If, on the other hand, the difference between gender-matching and mismatching conditions is unrelated to sluice projection, with RNP-contained reflexives not being constrained by regular binding conditions, or if the pied-piped PP is somehow ignored and does not serve to block a sluice continuation, then we would not expect to find an interaction of factors, i.e. the factor of Wh-Type should not have any effect on the size of this difference. In other words, we would predict that there should be a significant main effect of Gender Matching, with gender mismatching conditions being always significantly slower at (or just after) the reflexive compared to gender matching conditions, regardless of their compatibility with a sluice prediction.

4.3.2.2 Method

Following the original study, a word-by-word self-paced moving window display methodology was used (Just, Carpenter, & Woolley, 1982) (Linger© (http://tedlab.mit.edu/∼dr/Linger/) software). Participants sat in front of a computer monitor which displayed each sentence one at a time. All characters in the sentence were covered by dashes and separated by spaces denoting the spaces between words. Participants uncovered the sentence by pressing the spacebar[4]; each press simultaneously uncovered the next sentence region and covered the previous one in dashes, so that only one region was uncovered at any given time[5]. The time between button presses was recorded and used to infer processing complexity, with longer reaction times for a region indicating higher processing complexity. Aside from the experimental items, participants also saw 2.5 times as many filler items. All experimental items and half of the fillers were followed by comprehension questions to gauge participants' attention to sentence content. Stimuli were separated by a non-timed fixation cross at the centre of the screen. Participants could take as many breaks as necessary, provided this was done whilst the fixation cross was on-screen and before presentation of the next stimulus had been triggered.

4.3.2.3 Participants

Due to time constraints and a lack of available suitable subjects, to attain the same statistical power as in Yoshida et al. (2013), where 40 participants and 24 items had been used, we

[4]For greater RT accuracy, a Razer©Blackwidow gaming keyboard was used.

[5]Sentence regions were separated as in the original study, denoted in (4.2) by superscript numbers, with the first clause being presented all as one region and each word following that separately.

decided to instead use 24 participants and 40 items[6]. The results of twenty-four monolingual native Greek speakers of both Northern and Southern dialects are presented here (aged 19 - 45). All participants had completed high-school education (the majority had a university degree) and were healthy, right-handed, with no hearing or visual impairments (corrected eyesight allowed). Two additional subjects were originally tested, however their data was excluded as they responded with less than 80% accuracy to comprehension questions. All participants were compensated for their time with £6 for approximately 45 minutes of participation in the experiment[7].

4.3.2.4 Items

As mentioned above, in order to match the statistical power of the original study, a total of 40 experimental items, interleaved with 100 fillers, were randomised and presented in a Latin Square design to participants.

4.3.2.4.1 Experimental Stimuli

The experimental stimuli were designed based on the following pattern (129), as shown for instance in (130).

(129) Main Clause Subject (Male/Female family member (NOMINATIVE) + Female/Male Possessor proper name (GENITIVE)) + Main Verb + Indefinite Internal Argument (ACCUSATIVE) + PP (temporal/location specification), ‘but’/‘and’ + NEGATION + Embedding Verb(s) + Interrogative Pronoun (ACCUSATIVE) + Reflexive + PP + Embedded Verb + Embedded Subject + Continuation.

(130) a. I aderfi tou Chari parousiase pollous
Det..F.NOM *sister*.F.NOM *Det.*.M.GEN *Charis*.M.GEN *presented.3sg* *many*.M.ACC.PL.
pinakes stin ekthesi, alla den itan ksekatharo
paintings.M.ACC.PL. *at.Det.*F.ACC *exhibition*.F.ACC *but* NEG *was*.3SG *clear*.N.NOM
pjous pinakes tou eaftou tis apo
which.M.ACC.PL. *paintings*.M.ACC.PL. *Det.*.M.GEN *self*.M.GEN *she*.M.GEN*/hers* *from*
tin ekthesi charike to mondelo siopira.
*Det.*F.ACC *exhibition*F.ACC *was.pleased*.3SG *Det.*.N.NOM *model*.N.NOM *quietly*

[6] Potential issues which may have arisen from increasing the number of items in this way are addressed in this experiment’s Results section (4.3.3).

[7] This study, along with all subsequently listed studies in this chapter, was approved by the UCL Research Ethics committee under the Project ID Number *LING-2012-3*, with Dr Andrea Santi as primary investigator. All subjects read the UCL Information and Ethics form prior to beginning the experiment and gave informed consent for their data to be processed for the purposes of this study, expressly agreeing, furthermore, to participate knowing they were free to withdraw at any point. All data collected prior to May 2018 were handled according the the 1998 Data Protection Act, and all data collected subsequently were handled according the 2018 EU General Data Protection Regulations.

'Charis's sister presented many paintings at the exhibition, but it was not clear which paintings of herself the model was secretly pleased with.'

b. O aderfos tis Charas parousiase pollous
*Det.*M.NOM *brother.*M.NOM *Det.*F.GEN *Chara*F.GEN *presented.3sg* *many.*M.ACC.PL.
pinakes stin ekthesi, alla den itan ksekatharo
*paintings.*M.ACC.PL. *at.Det.*F.ACC *exhibition.*F.ACC *but* NEG *was.*3SG *clear.*N.NOM
pjous pinakes tou eaftou tis apo
*which.*M.ACC.PL. *paintings.*M.ACC.PL. *Det..*M.GEN *self.*M.GEN *she.*M.GEN*/hers* *from*
tin ekthesi charike to mondelo siopira.
*Det.*F.ACC *exhibition*F.ACC *was.pleased.*3SG *Det..*N.NOM *model.*N.NOM *quietly*
'Chara's brother presented many paintings at the exhibition, but it was not clear which paintings of herself the model was secretly pleased with.'

Stimuli were created to be as true to the original stimuli as possible. Specifically, with respect to the proper names, common, clearly gender-identifiable names were chosen; among the different conditions within each item, male and female names were further matched for length and number of syllables. With respect to the coordinator used, as per the original study, stimuli contained both the coordinators *and* ('ke') and *but* ('alla'). Furthermore, half of the stimuli used *how many* as the *wh*-element, with the other half using a *which*-NP; in all cases the 'correlate' NP was repeated as part of the 'remnant' (e.g. *which paintings* in (130)). Finally, half of the stimuli contained a feminine reflexive and the other half a masculine reflexive (equally divided between *which*-NP and *how many* stimuli).

Aside from these common characteristics with the original design, however, there were also certain points in which the Greek stimuli necessarily differed from the original due to language-inherent differences or in order to control for such differences. Specifically, in Greek, the reflexive is expressed using three words, literally translating to 'the self mine', with the DP appearing in singular Accusative and the possessive (additionally marked for gender in the 3^{rd} person) in Genitive. Furthermore, some of the representational nominals used require a PP complement and others an NP.GEN complement, hence the reflexive either appears as in (131) or (132), with the first pattern being the predominant one (appearing in 75% of items).

(131) pinakas tou eaftou tis
*painting.*M.NOM *Det.*M.GEN *self.*M.GEN *she.*F.GEN*/hers*
'painting of herself'

(132) scholiasmos jia ton eafto tis
*commentary*M.NOM *for Det.*M.ACC *self.*M.ACC *she.*F.GEN*/hers*

'commentary about herself'

Another difference to be noted is that Greek has a slightly different preferred word order for embedded clauses. The original target region pattern − excluding the conjunction and introductory VP (of varying lengths and content) − was: Wh-NP + reflexive + PP + **NP (Subject) + VP**. For Greek the embedded Subject and VP positions are reversed, resulting in the order Wh-NP + reflexive + PP + **VP + NP (Subject)**. This difference is likely to work in favour of any effects of structure projection by giving participants slightly longer to parse the potential sluice before disambiguation at the new Subject.

Furthermore, in the original design, the manipulation of sluice predictability relied on sluice non-compatible conditions containing P-pied-piping, whereas sluice compatible conditions contained P-stranding. Given P-stranding is not an overtly acceptable option in Greek, this effect needed to be replicated slightly differently. This effect was mimicked as closely as possible by choosing transitive verbs for the interrogative clause which can either pattern with a bare Accusative internal argument (133a, 134a) or a PP internal argument (133b, 134b). Hence, in the sluice compatible conditions the bare Accusative verb pattern would be used, whereas in the sluice non-compatible conditions the PP verb pattern would be used. There is some small change in meaning between the two options, similarly to how in English there is a slight difference between *which book he read* vs. *about which book he read*, however this does not have any effect on our manipulations since participants only see one condition per item.

(133) a. cherome kapjon
to.cherish someone.M.ACC
'to cherish someone'

b. cherome jia kapjon
to.be.happy for someone.M.ACC
'to be happy for/about someone'

(134) a. I aderfi tou Chari parousiase pollous
Det..F.NOM *sister*.F.NOM *Det.*.M.GEN *Charis*.M.GEN *presented.3sg many*.M.ACC.PL.
pinakes stin ekthesi, alla den itan ksekatharo
paintings.M.ACC.PL. *at.Det*.F.ACC *exhibition*.F.ACC *but* NEG *was*.3SG *clear*.N.NOM
pjous pinakes tou eaftou tis apo
which.M.ACC.PL. *paintings*.M.ACC.PL. *Det.*.M.GEN *self*.M.GEN *she*.M.GEN*/hers from*
tin ekthesi charike to mondelo siopira.
Det.F.ACC *exhibition*F.ACC *was.pleased*.3SG *Det.*.N.NOM *model*.N.NOM *quietly*

‘Charis’s sister presented many paintings at the exhibition, but it was not clear which paintings of herself the model secretly cherished.’

b. I aderfi tou Chari parousiase pollous
Det..F.NOM *sister*.F.NOM *Det.*.M.GEN *Charis*.M.GEN *presented.3sg* *many*.M.ACC.PL.
pinakes stin ekthesi, alla den itan ksekatharo
paintings.M.ACC.PL. *at.Det*.F.ACC *exhibition*.F.ACC *but* NEG *was*.3SG *clear*.N.NOM
jia pjous pinakes tou eaftou tis
for which.M.ACC.PL. *paintings*.M.ACC.PL. *Det.*.M.GEN *self*.M.GEN *she*.M.GEN/*hers*
apo tin ekthesi charike to mondelo siopira.
from Det.F.ACC *exhibition*F.ACC *was.pleased*.3SG *Det.*.N.NOM *model*.N.NOM *quietly*
‘Charis’s sister presented many paintings at the exhibition, but it was not clear which paintings of herself the model was secretly pleased with.’

With respect to the embedding verb, e.g. *is clear* in (134a), this was carefully chosen to not c-select for PP internal arguments so as to not create the temporary grammatical illusion in Wh-PP conditions that the *wh*-PP may be this verb’s internal argument (which would have counteracted the experimental manipulation). Similarly, the main clause verb (*presented*, here) was also carefully chosen to not c-select for the specific P used in the Wh-PP condition each time, so that this PP would indeed block a sluice projection. Finally, additional material of varying lengths was added at the end of the sentence as a buffer for sentence-wrap-up effects (Just & Carpenter, 1980).

4.3.2.4.2 Fillers

In addition to the experimental items, 100 filler items, i.e. 2.5 times as many as experimental stimuli, were included so that participants remained naïve as to the true purpose of the study. All fillers were matched in length with the experimental items, comprising of three clauses, and beginning with the same combination style of masculine family member plus feminine proper name (or vice versa) as the experimental items. The first clause was also presented all at once as one region. As there is no mention of what sort of sentences the original study used as fillers, these were equally divided among simple active and simple passive sentences, active and passive embedded questions, and a combination of active - passive and passive - active sentences.

4.3.2.4.3 Comprehension Questions

Comprehension questions targeting the antecedent subject or PP, embedded subject or continuation clause[8] followed all the experimental items and half of the fillers for a total of

[8] There was no indication of what sort of comprehension questions Yoshida et al. (2013) used.

82 comprehension questions. Half of these were false and half true.

4.3.3 Results

Before statistical analysis, the dataset was cleaned of non-representative data points, with all reaction times below $100ms$ and above $2500ms$ per word being removed as erroneous ($n = 4,308$, i.e. approximately 7% of all data points). Following the original study, the first presented region, comprising of the full 'antecedent' clause (marked with superscript '1' in Table 4.2), was excluded from the analysis as 'contextual' data, as were practice items and stimuli with incorrect comprehension question responses. Predicted reading times were then calculated per region per participant based on both experimental and filler items, by entering the number of characters per word and the cubic spline of the word's position in the sentence (Hofmeister, 2011) in a linear mixed effects model with log-transformed reaction times as the dependent variable and subject as a random effect, as shown in (4.1).

$$LogRT \sim Word.Position + Word.Length + (1|Subject) \tag{4.1}$$

Based on this model, residual reading times were then calculated per region per participant, by subtracting the actual RTs from the predicted RTs, with a negative value indicating the region was read faster than expected and was, therefore, plausibly easier to parse, and a positive value indicating the opposite, that the region was read more slowly than expected and was, therefore, harder to parse. Using residual RTs instead of raw RTs allows us to account for individual differences in reading speed, but also how RTs may differ with different word lengths and at different linear points in the sentence, irrespectively of the word's structural position (Ferreira & Clifton Jr, 1986); more importantly, however, using residual RTs instead of raw RTs helps protect against issues that are often inadvertently introduced when one begins to remove outlying data points, thereby disturbing otherwise perfectly minimal comparisons. For comparison, however, and given the original Yoshida et al. study reports raw RTs, raw RTs and their log-transformed counterparts are also reported in the Appendix (C.1.2). Furthermore, all analyses reported below were also run on raw and log-transformed raw RT data, with no difference in results. Following the original study analysis, data values more than 2.5*SDs from the mean as grouped by subject, condition and region were excluded from analysis (this was done using Tang's Linger Toolkit (Tang, 2014–2018)), leading to a total of 680 observations being dropped (approximately 1% of total observations).

Because different items varied slightly as to how many words were presented in each region of interest (see e.g. above for different RNPs in Greek requiring either PP or NP complements), several regions were collapsed across to make more representative larger regions, keeping different regions of interest separate. Table 4.3 shows how this was done.

Table 4.3: Greek SPR Region Collapse

Regions	Antecedent
Region Numbers	1
Example	*The sister the Charis presented many paintings at the exhibition*

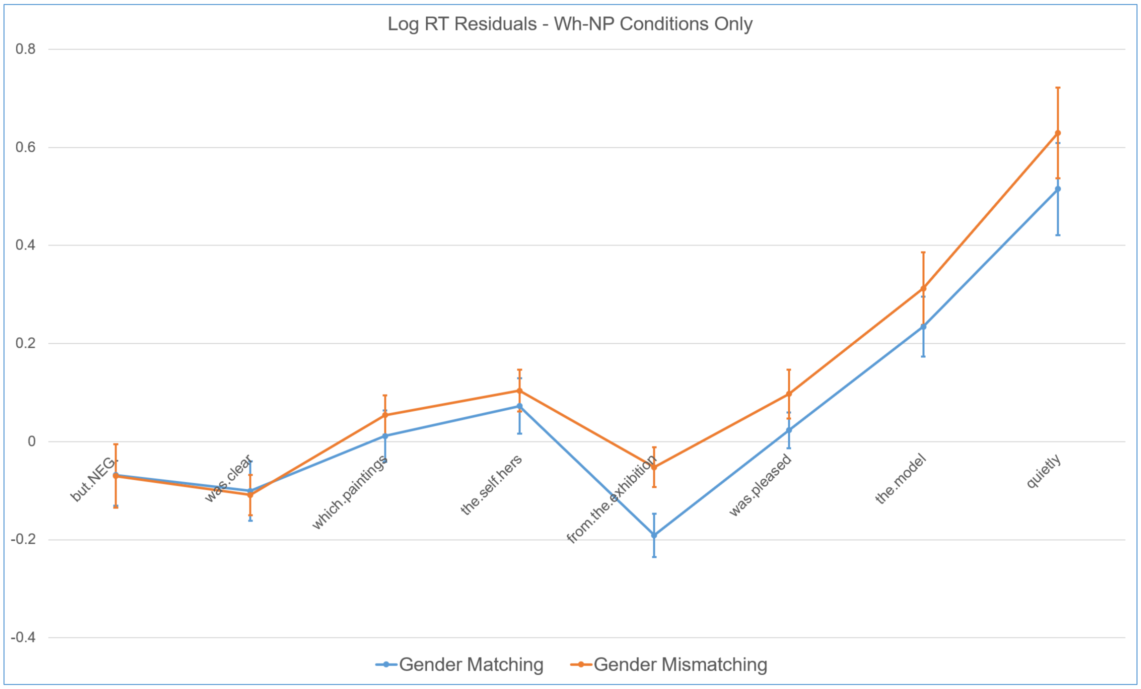
Regions	C	V	P	Wh	Reflexive	Embedded PP	Embedded V	Embedded NP	End
Region Numbers	2+3	4+5	6	7+8	9:12	13:15	16	17:19	20...
Example	*but NEG.*	*was clear*	*for*	*which paintings*	*the self hers*	*from the exhibition*	*was pleased*	*the model*	*quietly*

The Mean log10-transformed residual RTs plus standard error of mean (SEM) for these main regions of interest are shown in Figures 4.3a and 4.3b for Wh-NP and Wh-PP conditions respectively, with a common Figure in 4.4, for the item in (135).

(135) I aderfi tou Chari /O aderfos
Det..F.NOM *sister*.F.NOM *Det.*.M.GEN *Charis*.M.GEN *Det.*M.NOM *brother*.M.NOM
tis Charas parousiase pollous pinakes stin
*Det.*F.GEN *Chara*F.GEN *presented.3sg* *many*.M.ACC.PL. *paintings*.M.ACC.PL. *at.Det.*F.ACC
ekthesi, alla den itan ksekatharo (jia) pjous
exhibition.F.ACC *but* NEG *was*.3SG *clear*.N.NOM *for* *which*.M.ACC.PL.
pinakes tou eaftou tis apo tin
paintings.M.ACC.PL. *Det.*.M.GEN *self*.M.GEN *she*.M.GEN/*hers from Det.*F.ACC
ekthesi charike to mondelo siopira.
*exhibition*F.ACC *was.pleased*.3SG *Det.*.N.NOM *model*.N.NOM *quietly*
'Charis's sister presented many paintings at the exhibition, but it was not clear which paintings of herself the model secretly cherished.'

A maximal linear mixed effects model was successfully fitted to the data with log10-transformed residuals as dependent variable; the factors of Wh-Type (Wh-NP vs. Wh-PP) and Gender Matching (±) as fixed effects; and subjects and items as random effects, with random slopes and intercepts assumed for each. Interaction and main effects results for each region are shown in Tables 4.4 and 4.5, respectively, with the most interesting, significant ones highlighted here in order of appearance. Specifically, in the WH region (*which paintings*), a significant main effect of Wh-Type was found ($t > 2.40; p < .017$) with Wh-PP conditions being read on average faster than Wh-NP conditions (Mean Resid. LogRT difference: 0.10; Mean Resid. RT difference: 81.72ms; Mean LogRT difference: 0.01 ; Mean RT difference: 46.39ms). As expected, this difference was unaffected by the as-yet un-encountered reflexive gender, with no main effect of or interaction with the factor of Gender Matching. Moving on to the reflexive region itself, results showed no significant effects nor interactions, however there was a significant main effect of Gender Matching

(a) Greek SPR Results; Wh-NP Conditions



(b) Greek SPR Results; Wh-PP Conditions

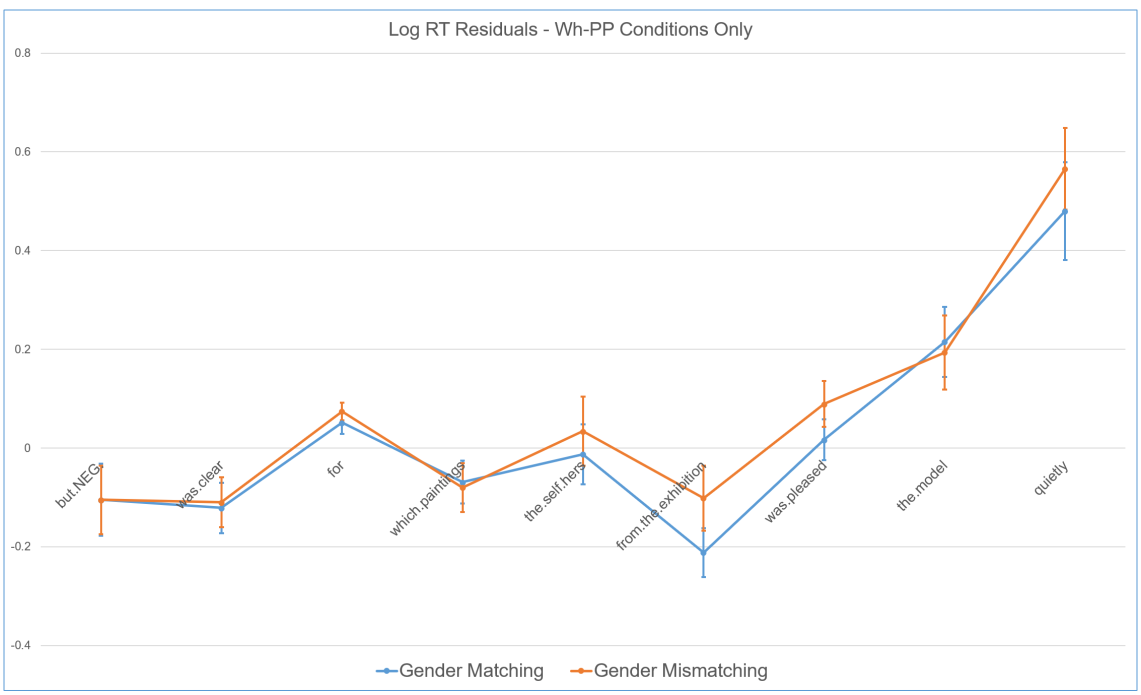
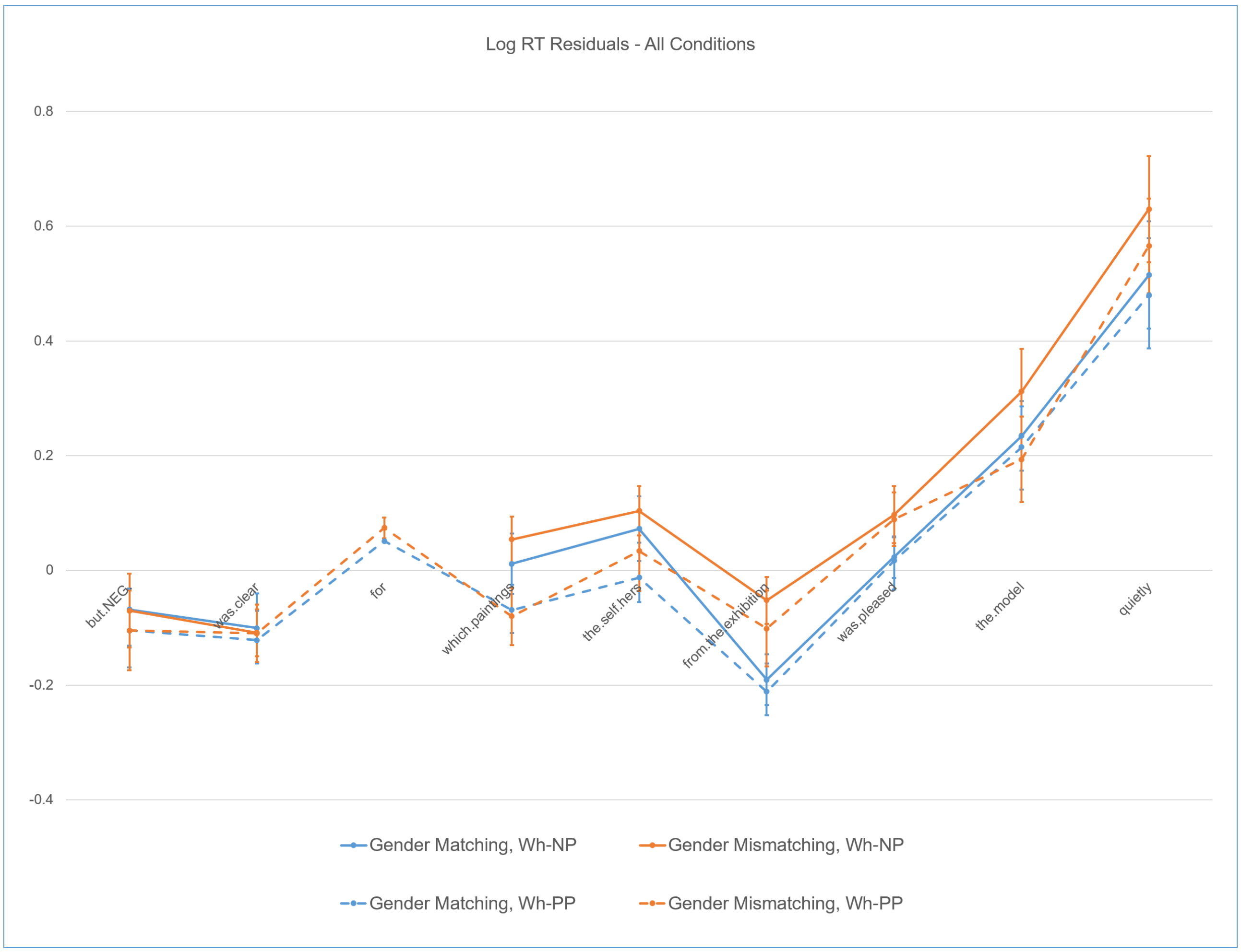


Figure 4.4: Greek SPR Log Transformed Results; All Conditions



at the immediately following PP region (*from the exhibition*) ($t > 2.31, p < .018$), with gender mismatching conditions being read significantly more slowly compared to gender matching conditions (Mean Residual Log RT difference: 0.13; Mean Residual RT difference: 79.45ms; Resid. LogRT Mean difference: 0.294; Mean RT difference: 87.762ms), the largest difference found between gender matching and mismatching conditions in any region. That having been said, this difference was unaffected by the factor of Wh-Type, i.e. there was no significant interaction between Wh-Type and Gender Matching either at or after the reflexive region ($t < .23; p > .82, n.s.$), nor was there a significant main effect of Wh-Type ($t < .59; p > .32, n.s.$), with no significant difference between Wh-NP and Wh-PP conditions. There were no further main effects nor significant interactions found at any of the following regions (i.e. embedded VP, embedded NP subject and sentence-final region). A full list of raw RT, log-transformed RT, residual and log-transformed residual Means with SEM per region per condition can be found in the Appendix (C.1.1).

Table 4.4: Greek SPR Analysis Results Summary by Region

	Statistical analysis results: Interactions			
	Wh-Type*Gender-Matching			
Region	β	SE	t	p
for (P)	N/A	N/A	N/A	N/A
which.paintings (Wh)	-0.061	0.075	-0.817	0.414
the.self.hers (Reflexive)	0.045	0.112	0.400	0.689
from.the.exhibition (Embedded PP)	-0.024	0.105	-0.228	0.819
was.pleased (Embedded VP)	-0.002	0.067	-0.026	0.979
the.model (Embedded NP)	-0.101	0.104	-0.972	0.331
quietly (End)	-0.013	0.120	-0.112	0.911

Table 4.5: Greek SPR Analysis Results Summary by Region (Continued)

	Statistical analysis results: Main Effects							
	Wh-Type				Gender-Matching			
Region	β	SE	t	p	β	SE	t	p
for (P)	N/A	N/A	N/A	N/A	0.024	0.035	0.700	0.484
which.paintings (Wh)	-0.092	0.038	-2.401	0.016*	0.017	0.037	0.461	0.645
the.self.hers (Reflexive)	-0.059	0.063	-0.935	0.350	0.035	0.067	0.523	0.601
from.the.exhibition (Embedded PP)	-0.030	0.051	-0.590	0.555	0.125	0.053	2.371	0.018*
was.pleased (Embedded VP)	-0.004	0.044	-0.082	0.935	0.063	0.039	1.610	0.108
the.model (Embedded NP)	-0.052	0.052	-0.988	0.323	0.025	0.052	0.477	0.633
quietly (End)	-0.038	0.056	-0.668	0.504	0.102	0.062	1.636	0.102

It should be pointed out here that a potential issue with decreasing the number of participants and conversely increasing the number of items that each participant saw compared to the original Yoshida et al. (2013) design is that this may have had a greater effect on participants' sensitivity here to the experimental manipulations of interest. In other words, as the experiment went on, participants may plausibly have become more de-sensitised to a potential gender mismatch effect driven by a sluice projection, particularly given the fact that no sluices were ever presented during the course of the experiment. Of course, by looking at the overall results of this study, one can see that even if this is indeed the case, there is still a significant difference between gender matching and gender mismatching conditions for apparently non-sluice related reasons; the question is whether this difference may have been additionally affected by the factor of Wh-Type, with Wh-NPs encouraging a larger mismatch effect compared to Wh-PPs, resulting in an interaction between the two factors, had there simply been fewer items. In order to check whether this could have been the case, a further exploratory analysis was run, including the factor of item presentation order as a fixed factor alongside Gender Matching and Wh-Type in a linear mixed effects model, comparing results from the first and second halves of the experiment for each subject. In this way, we were able to check whether Wh-Type had any effect on the difference between gender-matching and mismatching conditions in the first half of the experiment and whether this effect might have waned over the course of the experiment, leading to it not being detected in our main analysis above.

The linear mixed effects model including the fixed effects factors of Gender Matching ($\pm$), Wh-Type (Wh-NP vs. Wh-PP) and Order (1^{st} vs. 2^{nd} half), with log-transformed residuals[9] as the dependent variable found the same main effects pattern as in the original analysis, namely a main effect of Gender-Matching at the post-reflexive region ($t > 2.1, p < .048$), with a significant difference between gender-matching and mismatching conditions. In addition to this result, the model showed a significant main effect of the factor of Order in all regions, with subjects reacting significantly faster across the board throughout the second half of the experiment compared to the first half (all $t's > 7.48; p's < .0001$), something perhaps to be expected as participants became more familiar with the experiment format (see Table for all results). Importantly, however, there were no interactions of factors in any region (all $t's < 1.91; p's > .07$). This indicates that the difference found between gender-matching and mismatching conditions just after the reflexive region was not significantly affected by the factor of Order, nor the factor of Wh-Type in either the first or second halves of the experiment (see Appendix section C.1.3 for full analysis results). Given this difference between gender-matching and mismatching conditions did not change in any way over the course of the experiment, these results indicate that even if we had had a much smaller number of items, with a much smaller chance of subject desensitisation, we still would very

[9]The same results were found with non-transformed residuals, raw RTs and log-transformed RTs.

likely not have found a significant interaction between the factors of Gender Matching and Wh-Type at the region of interest (or any other region, for that matter) as Yoshida et al. (2013) did.

4.3.4 Discussion

As can be seen from our results above, it is quite clear that the original Yoshida et al. (2013) findings were not successfully replicated in Greek. Instead of finding an interaction between factors at (or just after) the reflexive region, a main effect of Gender Matching was found at the region immediately following the reflexive, namely the embedded PP region. This main effect was due to participants reading this region significantly more slowly in gender mismatching conditions compared to gender matching ones, regardless of Wh-Type. A main effect of Wh-Type was also found at the *wh*-XP region, with Wh-PP conditions being read significantly faster than Wh-NP conditions.

Beginning with the main effect of Gender Matching first, this very interesting result provides evidence for the hypothesis that the RNP-contained reflexive used in these two studies is not in fact constrained by the rules of binding theory, as has also been argued by others (Kaiser et al., 2005; Pollard & Sag, 1992; Reinhart & Reuland, 1993; Runner & Goldwater, 2011). If it were constrained by these rules, then we would have expected an interaction between the two factors, with a GMME occuring only in the condition where a fully hierarchical structure licensing this reflexive is predicted. Instead, we see a main effect of this factor, with the reflexive apparently generating an immediate search for its antecedent (Kazanina et al., 2007; Sturt, 2003; Van Gompel & Liversedge, 2003), regardless of the surrounding context. The reason we can see this effect in Greek, but not in English, is likely thanks to one of the main design advantages of the Greek study over the original, namely the lack of local attachment ambiguity very near the reflexive in the Wh-PP conditions. In the English study, even though adding this PP was meant to block a sluice projection by alerting participants to an as-yet-unseen verb, it was more likely that participants would instead interpret this PP in some alternative way, such as an AdvP (e.g. *but it was not clear for sure...*) or as the complement of the previous verb (e.g. *but we didn't remember to pay for the paintings...*). A predisposition for P-stranding over P-pied-piping in English may also contribute to such an alternative interpretation. When the parser realises this is not the correct interpretation, it is forced to recalibrate. This realisation will occur at or around the reflexive, meaning that this recalibration process may make the mismatching gender of the reflexive go unnoticed. In Greek, however, there is no such local attachment ambiguity, meaning that upon encountering the Wh-PP in these conditions, participants can clearly and swiftly interpret this as the argument of an as-yet-unseen verb. As such, they are not distracted by anything else when the reflexive is encountered, registering the gender mismatch in both Wh-NP and Wh-PP conditions equally. This result supports our hypothesis that the original Yoshida et al. (2013) study results may have been driven by this attachment ambiguity.

The main effect of Wh-Type found in the *wh*-region in the Greek results is also likely due to this lack of attachment ambiguity. Once the P is encountered in the Wh-PP conditions, the parser immediately recognises it as the beginning of an unseen verb's complement, and is thus able to quickly incorporate this P and its complement, the *wh*-XP. As the experiment goes on, again thanks to this lack of local attachment ambiguity, Greek participants may be more likely to begin to anticipate the study pattern, realising that a P encountered after a conjunction and negation is the beginning of a full PP complement. This means they may begin to predict a *wh*-XP after such Ps, or at the very least the P itself is a strong predictor of an upcoming DP, making these P complements less surprising and thereby easier to parse compared to when they are encountered without a preceding P to alert the parser[10]. That having been said, we did not find a significant interaction of Order with Wh-Type at the *wh*-region in our exploratory analysis, however this could be due to our looking very generally at the experiment in two halves, with this acclimatisation to the sentence pattern perhaps occurring sooner than the experiment's midway point. We leave this question for further exploration at a later point.

Despite our results, however, this does not mean to say that the Yoshida et al. (2013) conclusions were incorrect. It is still possible that we project ellipsis, in the form of sluicing, by default whenever given the opportunity, however the large GMME generated by the reflexive in the Greek study (coupled with confounding factors in the original study), may be masking the potentially subtler effects of such a projection. As the results currently stand, it is impossible to say for sure what the correct conclusion is. In order to distinguish between these two possibilities and try to answer the original question, one would need to remove both the reflexive and local attachment ambiguity from the equation and re-run the study. In order to address the original English results, we therefore decided to do just that and re-run a simplified study version in English[11].

4.4 Simplified English Self-Paced Reading Replication

4.4.1 Experimental Outline, Aim and Predictions

As concluded in the previous Discussion section, with the data at hand, it is currently impossible to distinguish whether we do indeed have a default tendency to project sluicing

[10]This also coincides with some of the Greek participants' comments after the study was over.

[11]It should be taken into consideration at this point that Yoshida et al. also had issues replicating their original findings in two follow-up study attempts (p.c. May, 2017). The researchers suspected that this was due, in part at least, to a dialectal difference, as the original study was conducted at Edinburgh University on a British English speaking set of subjects, whereas the follow-up studies were conducted at Northwestern University on a set of American English-speaking subjects. It is possible, however, that this replication problem was due to the original results being due to chance.

when presented with the opportunity to do so or not, much less be able to distinguish whether such a projection is also hierarchically organised or not. Given the main concern with both the original study and the Greek replication stemmed from the use of an RNP-contained reflexive, in order to answer this pressing question it is imperative that we find another way to measure the effects of such a projection, one which does not rely on the binding constraints of anaphors under sluicing. We therefore created a simplified version of the original study's design which took advantage of our inherent propensity for garden-pathing or slowing down during reading as a result of reanalysis. To directly compare the results of this study against those of the original, we also decided to run this follow-up in English.

The new stimuli follow the example in (136), a template of which is given in (137). The design is similar to the two previous studies in some aspects, i.e. in creating the context for a potential sluice projection and then foiling this projection. The first part of the sentence up until the comma serves as the antecedent for a potential sluice, by introducing an indefinite which can then be questioned. Following the comma, a sluice context is created with the contrastive coordinate *but* followed by a knowledge verb. The antecedent indefinite is questioned, but it is now followed by a new subject (*Annabelle*) and verb (*responded*), followed by an additional end clause. In other words, the main difference between these and the original stimuli is that, instead of the reflexive, the structure contains a new proper NP.

(136) Andrew heard some rumours at the pub, but we did not know for sure which rumours Annabelle responded to and I would like to find out somehow.

(137) Subject + Verb + Indefinite NPx (plural) + PP (locative or temporal adjunct), + BUT + it was not clear/we did not know (or other such phrase) + PP + which NPx (plural) + New Subject + New Verb + AND + I would like to find out somehow (or other such phrase).

As in the original and the Greek study, the factor of sluice predictability is manipulated by incorporating a pied-piped P directly before the *wh*-phrase or not, as shown in (138) (for consistency, I will use the term Wh-NP to refer to sluice compatible conditions and Wh-PP to sluice non-compatible conditions). The logic behind this design is similar to the original in that if we project sluices by default, then we should inherently prefer to create a sluice upon encountering the *wh*-phrase. To detect this projection, instead of relying on a GMME created by a reflexive, however, we are relying on a new sentence subject (i.e. *Annabelle*), which is encountered immediately following the *wh*-phrase. This subject serves to foil the sluice projection by introducing a new discourse referent and showing that the *wh*-phrase is actually part of an entirely new clause, not a sluice. If sluicing is indeed our most highly ranked or preferred continuation, then upon encountering this new subject which is inconsistent with the projected continuation, our results should show a surprisal slow-down at (or just after) this unexpected constituent. Furthermore, given a sluice should be expected in the Wh-NP condition, but not the Wh-PP one, we would anticipate a slow-down at this disambiguating subject only in the first condition. If, however, we find no significant difference between the two conditions, then this would indicate that a sluice prediction is not the most anticipated continuation.

(138) a. *Wh-NP Condition*
Andrew heard some rumours at the pub, but we did not know for sure which rumours Annabelle responded **to** and I would like to find out somehow.

b. *Wh-PP Condition*
Andrew heard some rumours at the pub, but we did not know for sure **to** which rumours Annabelle responded and I would like to find out somehow.

As mentioned previously, the original SPR design also contained the confounding element of a local attachment ambiguity in the Wh-PP conditions. To eliminate this possibility, or at the very least drastically reduce it, an adjunct PP (*for sure*) was introduced here between introductory verb (*know*) and *wh*-XP (*(to) which rumours*). This added adjunct PP reduces the likelihood that the pied-piped PP in the Wh-PP conditions could be anything other than the complement of a new, unseen verb. Another calibration compared to the original study is that the coordinating conjunction was limited to the contrastive *but*; this was due to the coordinators *and* and *but* being argued to lead to different syntactic expectations for upcoming material, such that *and* encourages more structural parallelism with the preceding clause compared to *but*, something which may in turn affect processing times for upcoming material (Lemke et al., 2018). Although there is not enough evidence yet for or against this idea, it was included here in order to err on the side of caution. Another important change to note is the addition here of an extra clause following the embedded verb (*... and I would like to find out somehow*), allowing us to account for potential spill-over effects from the embedded verb (e.g. due to closing of the *wh*-dependency), as well as sentence wrap-up effects.

In addition to the original study question regarding our potential tendency to project sluicing, this experiment also provided the opportunity to investigate another interesting claim. Earlier on, when discussing the results of Yoshida et al. (2013) in 4.2.2, we touched upon the status of P-pied-piping vs. P-stranding in English embedded clauses. Specifically, it has been argued that, as far as English is concerned, P-pied-piping is considerably less frequent compared to P-stranding, rendering it harder to parse, if not completely unacceptable (Cable & Harris, 2011; Heck, 2008). To the extent of our knowledge, there has not yet been an experimental study directly comparing the processing costs associated with incorporating such P-stranded vs. P-pied-piped constituents. The result of such an investigation would also provide more evidence as to whether it is indeed realistic to use a pied-piped P to indicate a different upcoming verb and foil a sluice projection in English, as intended in this and the original study designs.

To address this question here and examine how easy pied-piped PPs are to incorporate compared to P-stranded arguments, an additional factor was included in the experiment, that of antecedent verb type (P-less Verb vs. PP-Verb), manipulating whether a PP argument can be easily incorporated into preceding text or not. By substituting the antecedent verb for a verb which c-selects for a PP (e.g. *listen to*, (139)), this creates two additional, minimally different conditions (see Table 4.6).

(139) Andrew listened to some rumours at the pub, but we did not know for sure to which rumours Annabelle responded and I would like to find out somehow.

Table 4.6: Simplified English SPR Conditions

	Condition	Example
1	Wh-NP P-less Verb	Andrew heard some rumours at the pub, but I do not know for sure which rumours Annabelle responded to and I would like to find out.
2	Wh-PP P-less Verb	Andrew heard some rumours at the pub, but I do not know for sure to which rumours Annabelle responded and I would like to find out.
3	Wh-NP PP-Verb	Andrew listened to some rumours at the pub, but I do not know for sure which rumours Annabelle responded to and I would like to find out.
4	Wh-PP PP-Verb	Andrew listened to some rumours at the pub, but I do not know for sure to which rumours Annabelle responded and I would like to find out.

If the existing theoretical literature claiming that pied-piped PPs are generally unacceptable in English is correct, then we would expect to find that *wh*-phrases which are part of a pied-piped PP are always harder to parse compared to those which are P-stranded, regardless of how easily they can be incorporated into preceding text. In other words, the difference in reaction times between the Wh-PP and Wh-NP conditions at the *wh*-phrase (*which rumours*) should not be significantly affected by whether or not the antecedent verb can incorporate this P or not, i.e. there should not be a significant interaction between the factors of Wh-Type and Verb Type. If, on the other hand, the theoretical literature is wrong, and pied-piped PPs are not always unacceptable, then we can make two further predictions.

Firstly, if pied-piping is acceptable and we have some propensity towards a sluice or parallel structure prediction, then we would anticipate an interaction of Wh-Type and Verb Type such that a pied-piped PP is easier to incorporate for PP-Verbs, i.e. when the pied-piped PP can be incorporated into such a sluice prediction, compared to P-less Verbs, i.e. when it cannot. This would lead to a slow-down at the *wh*-phrase only in the sluice incompatible condition (Wh-PP, P-less Verb), i.e. we would expect a smaller difference in RTs at the *wh*-phrase between Wh-NP and Wh-PP conditions when paired with a PP-Verb compared to a P-less Verb. If this difference is indeed due to a sluice or parallel structure projection, then we would also anticipate a similar interaction at a later point in the sentence, namely the new subject (*Annabelle*) and/or verb (*responded*) when this sluice or parallel projection is foiled, respectively. Specifically, we would anticipate a slow-down at this sentence point for all the sluice-compatible conditions (Wh-NP, P-less Verb; Wh-NP, PP-Verb; Wh-PP, PP-Verb), but not the sluice incompatible condition (Wh-PP, P-less Verb). In other words, thanks to these potential projections being foiled at this sentence point, we would anticipate an interaction of factors in the same direction as previously, with a greater difference between Wh-NP and Wh-PP conditions when paired with a P-less Verb compared to a PP-Verb.

There is, however, also the possibility that we do not in fact have an inherent propensity to project a sluice or parallel structure when presented with the opportunity, leading to our second, slightly different prediction. In this case, as in the first prediction, when encountering the *wh*-phrase we may still anticipate an interaction of Wh-Type and Verb Type, such that pied-piped PPs are easier to incorporate for PP-Verbs rather than P-less Verbs, however in this case we predict this could simply be due to an effect of *priming*. In other words, simply having already seen the specific P in the antecedent in the PP-Verb condition may make it easier to parse when encountered a second time in the pied-piped PP, leading once again to a smaller difference at the *wh*-phrase between Wh-NP and Wh-PP conditions for PP-Verbs, where the P has been primed, compared to P-less Verbs, where it has not. If this is true, however, then this leads to a divergence in this second prediction compared to the first one with respect to later sentence stages. Specifically, since there is no prediction here of a sluice or parallel structure, we would not anticipate a slow-down at the later sentence points where such structures are foiled, namely the new subject (*Annabelle*) and verb (*responded*). In other words, we would not expect an interaction of factors at these points, with no difference predicted between any of the conditions.

So to recap, if pied-piped PPs are generally unacceptable, then we would expect to find a main effect of Wh-Type at the *wh*-phrase (*which rumours*), such that both Wh-PP conditions are always harder to parse compared to Wh-NP conditions, regardless of antecedent Verb Type. If, however, they are actually acceptable, then there are two possibilities. Firstly, if we project a sluice or parallel structure, this would make all three conditions compatible with this structure easier to parse compared to the one that is not compatible (Wh-PP, P-less Verb), leading to an interaction of factors at the *wh*-phrase; if this is true, then we would also anticipate the same direction of interaction at the later sentence stage when such a structure is foiled, namely the new subject (*Annabelle*) or verb (*responded*), with the non-sluice compatible condition (Wh-PP, P-less Verb) this time being easier to parse at this point than the other three. Secondly, if we do not project a sluice or parallel structure, but are instead affected by structural priming, then this would once again make the Wh-PP, PP-Verb condition where the pied-piped P has been primed, easier to parse than the Wh-PP, P-less Verb condition, where it has not, leading to an interaction of Wh-Type and Verb Type at this point. The difference with the first prediction is that in this case we would not anticipate any difference between conditions at the later sentence points where a sluice or parallel structure prediction are foiled, i.e. there should be no interaction of factors at the new subject or verb.

4.4.2 Methods

4.4.2.1 Experimental Design and Predictions

The two-level factors of Wh-Type (Wh-NP vs. Wh-PP) and Verb Type (P-less Verbs vs. PP-Verbs) were crossed to create the four conditions presented in Table 4.7 (repeated from above with added region numbers).

Table 4.7: Simplified English SPR Conditions (annotated)

	Condition	Example
1	Wh-NP, P-less Verb	Andrew heard some rumours at the pub,[1] but[2] I[3] do[4] not[5] know[6] for[7] sure[8] which[9] rumours[10] Annabelle[11] responded[12] to[13] and[14] I[15] would[16] like[17] to[18] find[19] out.[20]
2	Wh-PP, P-less Verb	Andrew heard some rumours at the pub,[1] but[2] I[3] do[4] not[5] know[6] for[7] sure[8] to[9] which[10] rumours[11] Annabelle[12] responded[13] and[14] I[15] would[16] like[17] to[18] find[19] out.[20]
3	Wh-NP, PP-Verb	Andrew listened to some rumours at the pub,[1] but[2] I[3] do[4] not[5] know[6] for[7] sure[8] which[9] rumours[10] Annabelle[11] responded[12] to[13] and[14] I[15] would[16] like[17] to[18] find[19] out.[20]
4	Wh-PP, PP-Verb	Andrew listened to some rumours at the pub,[1] but[2] I[3] do[4] not[5] know[6] for[7] sure[8] to[9] which[10] rumours[11] Annabelle[12] responded[13] and[14] I[15] would[16] like[17] to[18] find[19] out.[20]

Explanatory Note: Superscript numbers denote presented regions (with the antecedent being presented all at once).

When discussing the possibility of a sluice or parallel structure projection, based on the original study logic, out of these four conditions, only the Wh-PP, P-less Verb condition is incompatible with a sluice projection at the *wh*-phrase (**heard to*). The pied-piped PP in this condition is also not structurally primed by the presence of the same P alongside the antecedent verb, as opposed to the PP in the Wh-PP, PP-Verb condition. As made clear in the previous section, it is the difference in participants' reading behaviour at several sentence points for this condition in particular, therefore, compared to the other three conditions that we are most interested in. Specifically, the sentence points of interest are the *wh*-phrase *which rumours*, i.e. whether the Wh-PP, P-less Verb condition is significantly slower than the rest here or whether both Wh-PP conditions are slower than Wh-NP conditions; and the new subject (*Annabelle*) and verb (*responded*), i.e. whether the Wh-PP, P-less Verb condition is significantly faster than the rest here, or whether there is no difference between any conditions.

Since one of the key design aspects of the study is the introduction of a new, embedded subject (*Annabelle*) and the effects we are looking for do rest upon participants actually realising that this is a different person to the antecedent subject (*Andrew*), in order to maximise discourse prominence (Sanford et al., 1988) and thereby also the potential for a slow-down in RTs, we decided to use proper NPs, with different genders, instead of common NPs as the antecedent and embedded subjects. That having been said, it should be noted that the introduction of a new discourse entity may in itself also result in a slow-down at the embedded subject, *Annabelle*. If the results do indeed show a slow-down at *Annabelle* that is due solely to introducing a new discourse referent and not the disambiguation of a sluice projection, then we would expect to find this slow-down across all four conditions, i.e. there should be no interaction here.

The second main design aspect of the study is whether or not the pied-piped PP can be easily integrated into preceding text or not, based on whether the antecedent verb is a P-less (*hear*) or PP-Verb (*listen*). To allow for an exploratory analysis of whether the type of P may have any effect on the ease of integrating P-pied-piping, perhaps due to the sheer frequency of each P, equal numbers of four of the most frequent English prepositions were used: *for*, *with*, *at*, *to* (8 of each). With respect to sheer frequency of occurence, based on the SUBTLEX-UK corpus (Van Heuven et al., 2014), these can be ordered from most to least frequent as: *to* (LogFrequency (Zipf[12]) value: 7.42), *for* (LogFrequency (Zipf) value: 6.93), *with* (LogFrequency (Zipf) value: 6.75) and *at* (LogFrequency (Zipf) value: 6.69), with this ordering also reflecting the contextual diversity of the Ps (i.e. in the context of the SUBTLEX-UK corpus, the number of BBC films and series the word appears in), from

[12] This scale was originally proposed by Van Heuven et al. (2014) as a measure of absolute lexical frequency. The scale is logarithmic and calculated as:

$$Zipf = log10(\frac{frequency.count + 1}{corpus.size + frequency.counts}) + 3 \tag{4.2}$$

As a scale, it overcomes two important issues which handicap the usual measure of lexical frequency, i.e. frequency per million : a) the scale allows calculations for words occurring $< 1pm$, quite a common occurrence; b) it assigns a meaningful absolute frequency rating which is independent of size of the original corpus the word originated in, i.e. a word with a Zipf frequency of 2 is always considered a low frequency word, whereas one of 5 and above is a high frequency word.

most to least diverse (Brysbaert, Mandera, & Keuleers, 2018). If parsing pied-piped Ps is in any way linked to this frequency or contextual diversity, then we may anticipate if not a difference between all Ps, then perhaps a difference between the most commonly occurring one (*to*) and the least commonly occurring one (*at*).

In order to overcome one final limitation of the original study, in that it never presented participants with a sluice or parallel continuation in the fillers or experimental items, rendering it quite likely that participants may have eventually stopped projecting such continuations, a large number of sluices were included in the fillers (50% of fillers) here.

4.4.2.2 Method

As in both the original and Greek SPR studies, a word-by-word self-paced moving window display method (Just et al., 1982) was used, utilising Linger© (http://tedlab.mit.edu/∼dr/Linger/), and conducted at the UCL Chandler House Linguistics lab facilities. Participants sat in front of a computer screen which displayed each sentence one at a time. Dashes covered all the characters and spaces denoting the spaces between words separated them. Participants pressed the spacebar[13] to uncover the sentence one region at a time (see Table 4.7 for sentence regions), with each press simultaneously uncovering the next region and covering the previous one in dashes, so that only one region was uncovered at any given time. To monitor participants' attention to the task at hand, all experimental items and 50% of fillers were followed by a comprehension question (see below for more details). A non-timed fixation cross at the centre of the screen separated stimuli, with participants being allowed to take as many breaks as necessary whilst the fixation cross was on-screen and before triggering the presentation of the next stimulus.

4.4.2.3 Participants

Using UCL Psychology's online SONA Subject Pool (https://uclpsychology.sona-systems.com/), we recruited 40 subjects from a range of backgrounds (UCL/other university students - older working professionals). Participants were healthy, right-handed, monolingual, native British English speakers, born and raised in the UK, with no hearing or visual impairments (corrected eyesight allowed) and no history of learning difficulties, neurological or psychiatric issues. All participants responded to a pre-screening questionnaire, consented to taking part in the experiment beforehand and were compensated with £4 for approximately 20 - 30 minutes of participation.

[13]As in the Greek SPR study, for greater RT accuracy, a Razer©Blackwidow gaming keyboard was used.

4.4.2.4 Stimuli

In total, 32 experimental stimuli and 64 filler items were presented to participants in a Latin Square design utilising Linger's inbuilt pseudo-randomisation system so that equal numbers of each condition were obtained based on a pre-defined total number of participants. To gauge participants' attention and understanding, half the fillers and all experimental stimuli were followed by a comprehension question, equally split between questions targeting a) the antecedent subject; b) the embedded subject; c) the antecedent PP; and d) the 'extra' sentence-final CP. Out of these, half were true and half false.

The experimental stimuli were all approximately the same length in words (Min: 25; Max: 28; Mean: 26.65; SD: .92). Fillers approximately matched the experimental stimuli in sentence length (Mean: 23.5; SD: 2.2) and consisted of 50% sluices (140), 25% relative clauses (141), 12.5% simple actives (142) and 12.5% simple passives (143). Half of each category began with a male name and half with a female name. A large percentage of sluices was used in the fillers to stop participants from becoming de-sensitised to the possibility of encountering this structure, given that all experimental stimuli eventually foiled a sluice projection. All fillers, furthermore, matched the experimental stimuli antecedent pattern as closely as possible so participants would remain naive as to the study's true target pattern. A full list of all stimuli are given in the Appendix (C.2.3).

(140) Roger told us some stories at the dinner, but I do not remember too well which stories and I would like to find out somehow.

(141) Joshua heard some rumours at the dinner party, which his mother hosted and couldn't wait to tell us about them.

(142) Ivy tidied up the house in the afternoon, but it was her amazing cooking that really impressed her family and friends.

(143) Celia was finally made a partner at the law firm, but she cared more about the better hours than the increase in salary.

4.4.3 Results

The dataset was cleaned of all extreme, non-representative data points, with reading times below $100ms$ and above $2500ms$ being removed as erroneous ($n = 3,200$, i.e. approximately 4% of the total number of data points). Following the Yoshida et al. (2013) and Greek SPR studies, the first presented region, comprising of the full 'antecedent' clause (marked with superscript '1' in Table 4.7), was excluded from the analysis, as were practice items and stimuli with incorrect comprehension question responses. To account for individual participants' reading speed, as well as how different word lengths and different linear positions in the sentence (irrespectively of the word's structural position) may affect reading times beyond our manipulations (Ferreira & Clifton Jr, 1986) and for the same reasons as in the Greek SPR study, all our models used residual reading times. Based on experimental and filler items, predicted reading times were calculated per region per participant, based on the number of characters per word and the word's position in the sentence. This was done by including number of characters and cubic spline of word position as fixed effects in a linear mixed effects model with raw or log-transformed reaction times as the dependent variable and subject as random effect (4.3).

$$LogRT \sim Word.Position + Word.Length + (1|Subject) \tag{4.3}$$

Residual reading times were then calculated per region per participant, by subtracting the observed from the predicted RTs, with negative values indicating faster reading times than predicted and positive values indicating the opposite[14]. Data values more than 2.5 SDs from the mean (grouped by subject, condition and region) were further excluded from analysis using Tang's Linger Toolkit (Tang, 2014–2018) ($n = 922$, i.e. approximately 1% of total data points).

Although different items did not vary greatly as to how many words were presented, some adjacent words were collapsed across to make more representative regions, as in the Greek study, though to a lesser extent and always keeping different regions of interest separate. Specifically, the regions collapsed across were the AdvP (*for sure*), the Wh-NP (*which rumours*), the embedded verb if it contained an auxiliary, and the final extra VP (*would like to find out...*).

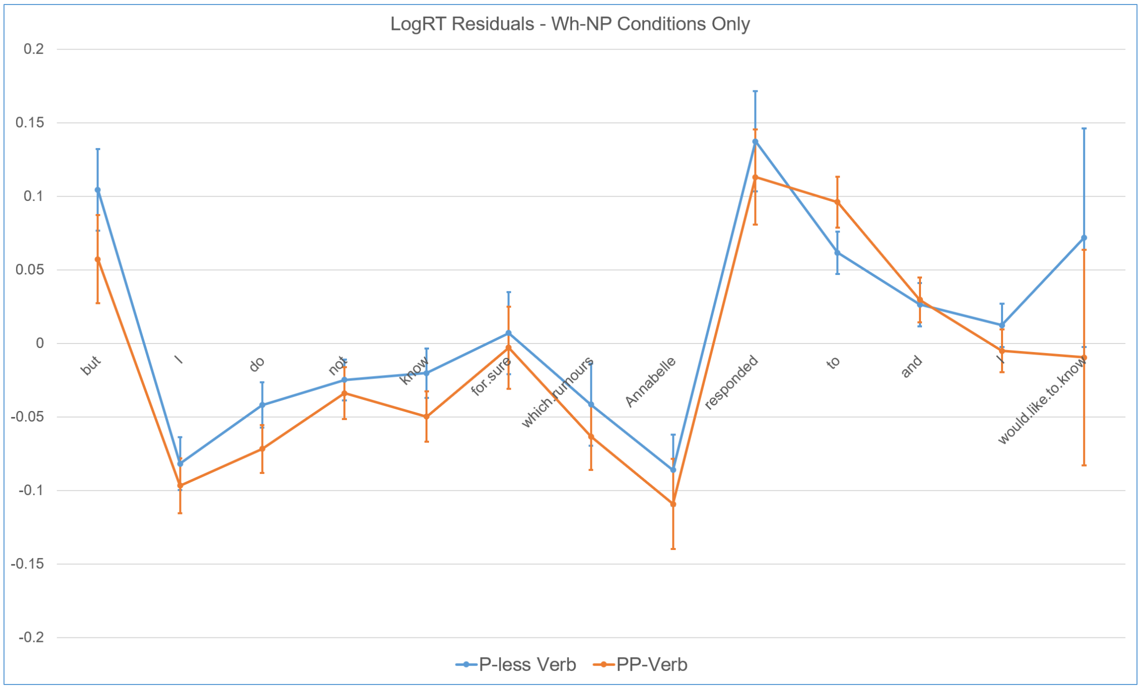
Collapsing across subjects, the Mean log10-transformed residual RTs (with SEM error bars) for each region of interest are given for all conditions together in graph 4.6, with separate graphs for the two Wh-NP conditions and Wh-PP conditions in Figures 4.5a and 4.5b, respectively[15].

A linear mixed effects model with log-transformed residuals as the dependent variable, the factors of Wh-Type (Wh-NP vs. Wh-PP) and Verb Type (P-less Verb vs. PP-Verb) as fixed effects and subjects and items as random effects (with random slopes and intercepts assumed) was successfully fitted to the data. Main effects and interactions for each region of interest separately are shown in Tables 4.8 and 4.9, respectively. With 32 experimental items and 40 subjects, there were a total of $1,280$ observations per region (aside from the P regions, which were analysed separately according to their sentence position, i.e. separately for the two Wh-NP conditions vs. the two Wh-PP, with 640 observations each).

Results showed a significant main effect of Wh-Type at a number of regions, a significant

[14] Although I believe that residual RTs are more accurate in their representation of reading effects compared to raw RTs, given the original Yoshida et al. (2013) study reports raw RTs, these, along with their log-10-transformed versions are also presented here in the Appendix (C.2.2). Furthermore, as in the Greek study, all the below analyses were also conducted on both raw and log-10-transformed raw RTs, with no differences

(a) English SPR Results; Wh-NP Conditions



(b) English SPR Results; Wh-PP Conditions

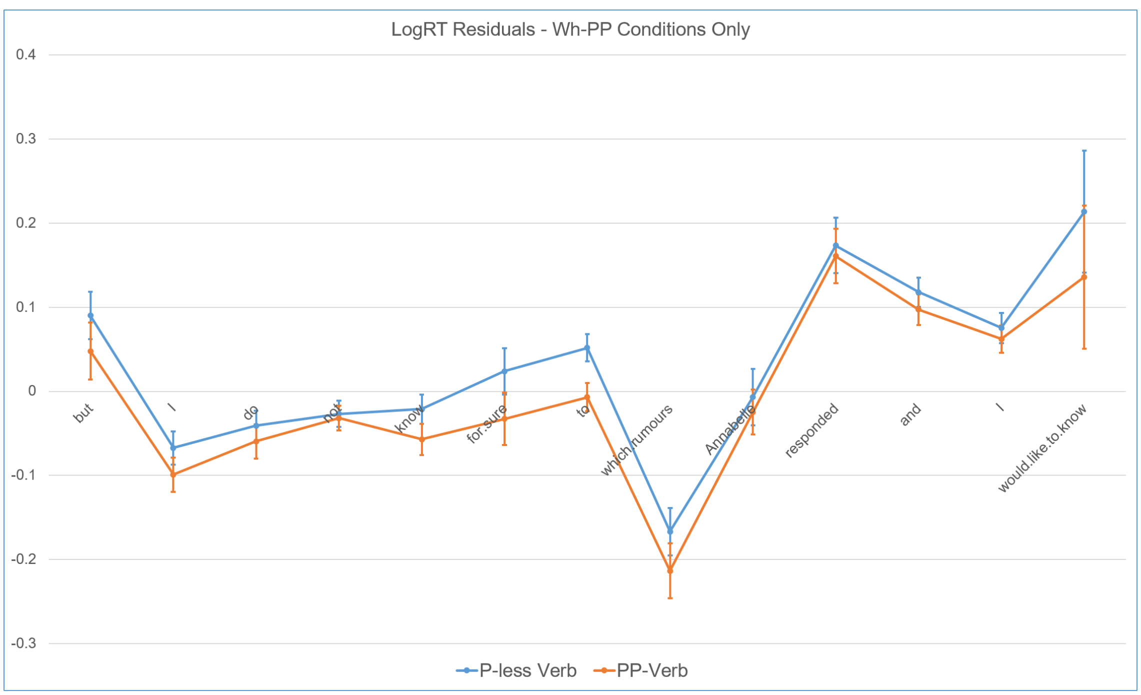


Figure 4.6: English SPR Results; All Conditions

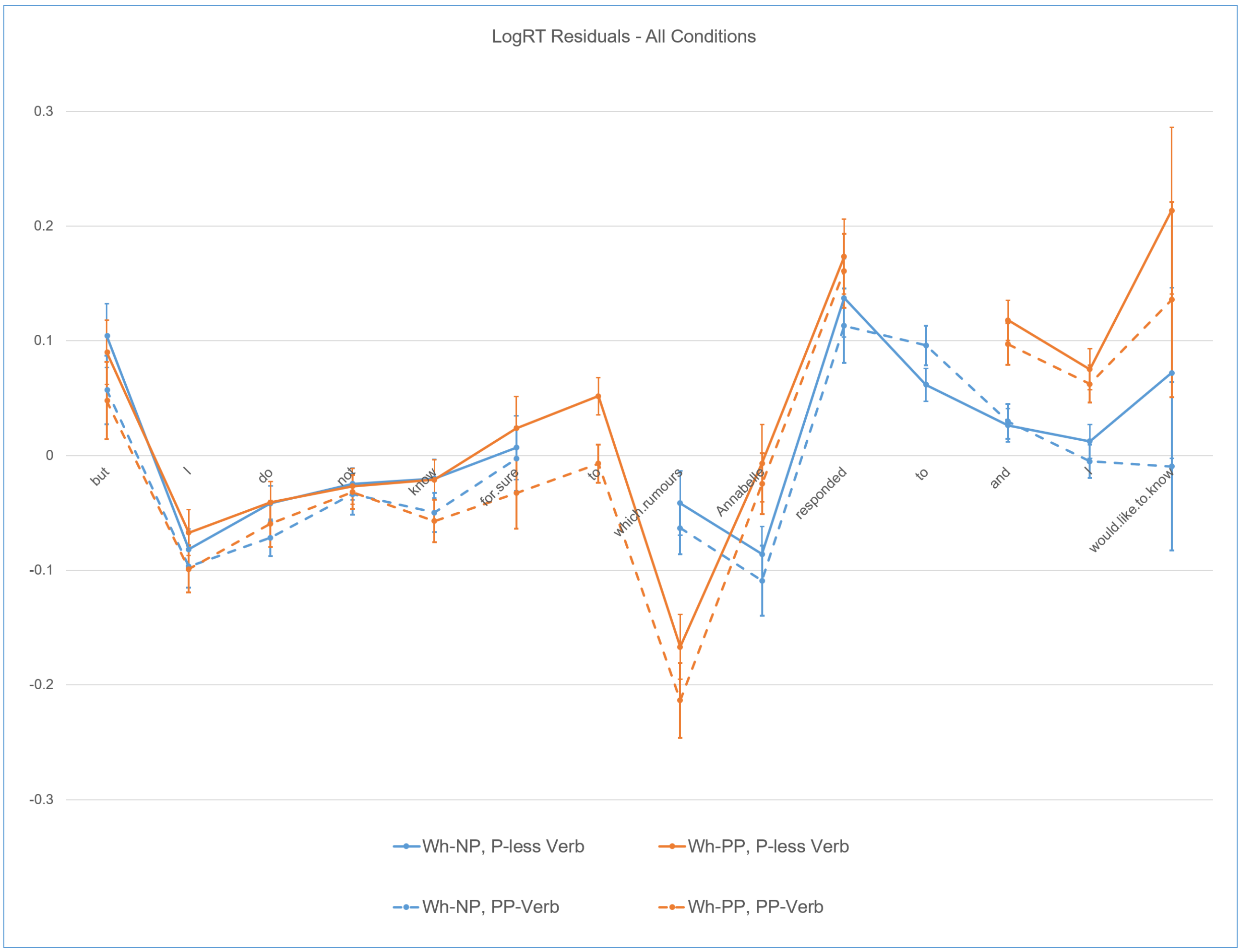


Table 4.8: English SPR Analysis Results Summary by Region (Main Effects)

	Statistical analysis results: Main Effects							
	Wh-Type				Verb Type			
Region	β	SE	t	p	β	SE	t	p
to (P; Pied-Piped)	N/A	N/A	N/A	N/A	-0.059	0.026	-2.238	0.025*
which.rumours (Wh)	-0.138	0.031	-4.396	0.000***	-0.034	0.029	-1.197	0.231
Annabelle (New Subject)	0.082	0.020	4.157	0.000***	-0.022	0.021	-1.043	0.297
responded (New Verb)	0.042	0.024	1.747	0.081	-0.018	0.028	-0.644	0.520
to (P; Stranded)	N/A	N/A	N/A	N/A	0.034	0.022	1.545	0.122
and (C2)	0.079	0.021	3.706	0.000***	-0.008	0.017	-0.497	0.619
I (Subject 3)	0.065	0.016	4.060	0.000***	-0.015	0.015	-0.993	0.321
would.like.to.know (Sentence End)	0.144	0.056	2.544	0.011*	-0.080	0.056	-1.425	0.154

Table 4.9: English SPR Analysis Results Summary by Region (Interactions)

	Statistical analysis results: Interactions			
	Wh-Type*Verb Type			
Region	β	SE	t	p
to (P; Pied-Piped)	N/A	N/A	N/A	N/A
which.rumours (Wh)	-0.025	0.064	-0.394	0.694
Annabelle (New Subject)	0.006	0.045	0.126	0.900
responded (New Verb)	0.011	0.055	0.205	0.837
to (P; Stranded)	N/A	N/A	N/A	N/A
and (C2)	-0.024	0.030	-0.787	0.431
I (Subject 3)	0.005	0.032	0.140	0.889
would.like.to.know (Sentence End)	0.004	0.107	0.037	0.971

main effect of Verb Type at a single region and no significant interaction between the two factors at any region (all t's$< .79$; all p's$> .431$). Specifically, a significant main effect of Wh-Type was found at the first region of interest, the *wh*-phrase (*which rumours*) ($t > 4.41; p < .00001$), with Wh-PP conditions (Resid. LogRT Mean: $-0.19ms$; Resid. RT Mean: $-84.38ms$) being read on average faster than Wh-NP conditions (Resid. LogRT Mean: $-0.05ms$; Resid. RT Mean: $-21.02ms$; Resid. LogRT Mean difference: $0.14ms$; Resid. RT Mean difference: $63.35ms$), regardless of Verb Type. At the next region of interest, the new subject (*Annabelle*), a significant main effect of Wh-Type was also found ($t > 4.27; p < .00002$), this time with Wh-PP conditions (Resid. LogRT Mean: $-0.02ms$; Raw RT Mean: $-84.38ms$) being read on average more slowly than their Wh-NP counterparts (Resid. LogRT Mean: $-0.10ms$; Resid. RT Mean: $2.71ms$; Resid. LogRT Mean difference: $0.08ms$; Resid. RT Mean difference: $37.73ms$), again regardless of Verb Type. The difference between Wh-PP and Wh-NP conditions was then minimised at the next region of interest, the new verb (*responded*) ($t < 1.58; p > .114, n.s.$), only to become significant again at the complementizer *and* ($t > 4.64; p < .001$), with Wh-PP conditions (Resid. LogRT Mean: $0.11ms$; Resid. RT Mean: $45.65ms$) being read on average more slowly than the Wh-NP ones (Resid. LogRT Mean: $0.03ms$; Resid. RT Mean: $11.02ms$; Resid. LogRT Mean difference: $0.08ms$; Resid. RT Mean difference: $34.63ms$), once again irrespectively of Verb Type. This significant difference between Wh-PP and Wh-NP conditions in the form of a main effect carries through to the third sentence subject (I) ($t > 4.06; p < 0.001$) with Wh-PP conditions (Resid. LogRT Mean: $0.07ms$; Resid. RT Mean difference: $27.68ms$) again being read more slowly than Wh-NP conditions (Log RT Mean: $0.004ms$; Resid. RT Mean: $3.82ms$; Resid. LogRT Mean difference: $0.07ms$; Resid. RT Mean difference: $23.86ms$).

A significant main effect of Verb Type was found at the pied-piped P region (*to*) ($t > 2.24; p < .025$) with P-less Verb conditions being read significantly more slowly than PP-Verb conditions (P-Less Verb: Resid. LogRT Mean: $0.05ms$; Resid. RT Mean: $27.27ms$; PP-Verb Resid. LogRT Mean: $-0.01ms$; Resid. RT Mean: $1.14ms$; Resid. LogRT Mean difference: $0.02ms$; Resid. RT Mean difference: $14.21ms$). Obviously this region cannot show an interaction as it is only present in the Wh-PP conditions, however there was no significant difference between P-less and PP-Verbs at the stranded P region ($t < 1.545; p > .122, n.s.$).

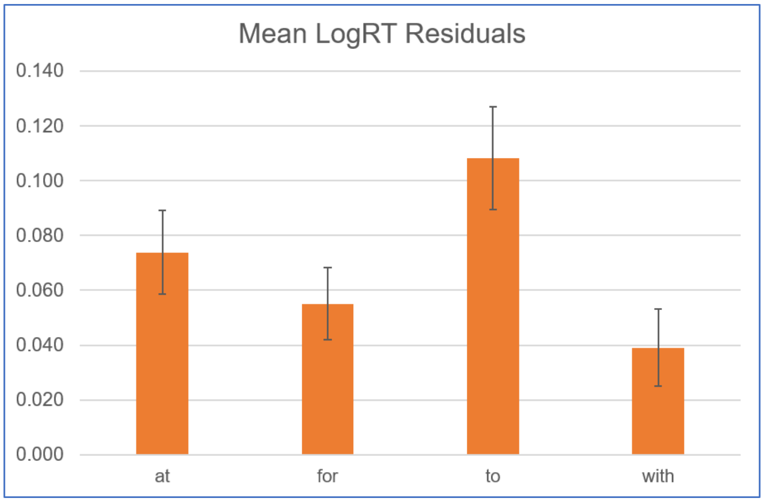
In an exploratory analysis, the effect of P type was also investigated at the P regions (pied-piped and stranded) as well as their immediately following regions, namely *wh*-phrase, new verb and second complementizer (*and*). In order to do so, the four-level factor of P Type (*to* vs. *for* vs. *with* vs. at) was added to the LME model using R's built-in deviation coding (comparing each level to the grand mean) to see if it contributed to the model's fit above and beyond the factor of Verb Type when examining only pied-piping conditions.

found in results.

[15]Numerical tables with these plus raw residual RT means and SEMs are available in the Appendix.

The factor had almost no effect on the model's fit at any region ($t's < 1.8; p's > .06$), apart from the second complementizer, where one level of P type, namely *with*, was significantly faster at being incorporated compared to the other factor levels ($t > 2.24; p < .025$) (4.7).

Figure 4.7: PP Type Means with SEM for C2(*and*); Pied-Piped Conditions only



4.4.4 Discussion

With respect to the study's original question, i.e. whether there is potentially an effect of sluicing prediction at the disambiguating subject *Annabelle* or the directly following verb region (*responded*), the fact that the one condition which was incompatible with a sluice projection, the Wh-PP, P-less Verb condition, did not behave any differently to the other conditions at either of these regions (or indeed any region at all), indicates this is not the case. This lack of significant interaction between factors lends further support to our hypothesis that the original Yoshida et al. study results were likely driven by properties of the reflexive itself and not some inherent tendency to predict sluices. Instead, here we found a series of interesting main effects at different regions. Starting from the beginning of the sentence, let us examine each of these effects in turn.

At the pied-piped P, we see a significant difference between the P-less Verb and the PP-Verb conditions, with the latter being read much faster than the former. This would indicate that it is significantly easier to incorporate a pied-piped P when it is compatible with the preceding text. To check whether this advantage is due to a sluice prediction or simply structural priming, we must seek evidence later on in the structure, namely at the new subject (*Annabelle*) and/or new verb (*responded*). The lack of main effect for Verb Type or interaction between factors at these regions, however, indicates that integrating the pied-piped P more easily with a PP-Verb is more likely to be due to priming than to predicting a sluice or parallel structure, given there is no surprisal slow-down at the point where such a structure would be foiled.

Moving on to the *wh*-phrase itself (*which rumours*) we see a significant difference here between Wh-NP and Wh-PP conditions, with the latter being read much faster than the former. This is similar to what we found at the same region in the Greek SPR study. The key difference between these two studies and the original Yoshida et al. study is that here there is no local attachment ambiguity at the pied-piped P. This means that participants can quickly and effectively identify the pied-piped P as the start of an unseen verb's complement[16]. Having made this identification, they can furthermore anticipate a P-complement (*which rumours*), speeding through it and knowing to keep the full PP active in mind until the open dependency is resolved, i.e. they encounter the new verb. Despite our best efforts to keep participants naïve as to our experimental manipulations, a coding error meant that all the sluice stimuli participants were exposed to in the fillers did *not* contain a pied-piped P. This means that participants may, over the course of the experiment, also have become aware that a pied-piped P is always followed by a *wh*-phrase, making it easier to anticipate and parse. The speed-up at the *wh*-phrase could, therefore, be due to participants expecting a *wh*-phrase after the pied-piped P, i.e. the predictability of *wh*-XP as w_i given all words up to w_{i-1} rising over the course of the experiment (J. Hale, 2001; Levy, 2008a; Smith & Levy, 2013).

Interestingly, moving on to the new subject, *Annabelle*, we again see a main effect of Wh-Type, but this time with Wh-PP conditions being read significantly more slowly than their Wh-NP counterparts. This may indicate that keeping active an open dependency involving a pied-piped PP is harder than one with a bare argument; this, in turn, may make it harder to integrate a new discourse referent (Gibson, 1991, 1998) or simultaneously create an additional open dependency whilst keeping this PP-dependency active. Regardless of this speculation, however, the key fact that we do not see a significant interaction of factors at this or the next region, where a sluice or parallel continuation are disambiguated, provides an answer to our original question regarding the default predictability of such structures, lending further evidence to their not being automatically projected here. In other words, the one condition which is not compatible with a sluice (Wh-PP, P-less Verb) is not significantly faster at any of these regions, a result we might expect had sluicing been predicted.

Moving forward in the sentence, we see no significant results at the new verb (*responded*), with all conditions being read at approximately equal speed. There is also no significant main effect of Verb Type at this region, the stranded P region or any following regions, as mentioned earlier. Finally, there are significant main effects of Wh-Type at the next three regions (the second complementizer *and*, new clause subject *I* and end of sentence *would like to know...*), with Wh-PP conditions being read significantly more slowly than

[16]The fact that, as we saw above, we do not find evidence for sluicing prediction at the new subject or verb also lends support to this interpretation.

their Wh-NP counterparts, regardless of Verb Type. Altogether, this lack of main effects and interaction at the new verb and stranded P, along with the significant main effects following, show that participants' behaviour at these sentence points is not driven by the projection of a sluice or parallel structure, nor is it affected by whether the P in question was seen in the antecedent or not. If we were to speculate on the origin of this behavioural pattern, we would tentatively suggest that it is likely due to it being significantly harder to close an active dependency involving a pied-piped P compared to one involving a stranded P.

Finally, there appears to be no significant advantage conveyed to parsing of pied-piped or stranded Ps nor their arguments based on the frequency of the P in question or its contextual diversity.

In conclusion, with respect to the original study's question, whether we actively predict sluicing when presented with the opportunity, neither this nor the Greek study provide any evidence in favour of a positive answer. There is also no evidence that we actively project a parallel continuation. With respect to the ease or not of parsing a pied-piped PP and whether this is affected by how easily it can be integrated into the preceding text, it appears significantly easier to parse the pied-piped P itself when the antecedent contains the same P, however it is not significantly harder to parse the fronted *wh*-phrase when it is preceded by a P, indeed the opposite appears true. As such this does not appear to provide direct evidence in favour of the claims made by Heck (2008) and Cable and Harris (2011) that pied-piped PPs are completely unacceptable. The only tentative finding which could potentially be in favour of these claims is that it appears to require more cognitive effort to close an open dependency involving a pied-piped PP vs. a stranded one, however this finding requires further investigation to be confirmed.

At the end of these experiments, we have not gathered any evidence to suggest that the parser predicts sluicing or parallel continuations as some form of default strategy. We have some evidence for primed pied-piped Ps being easier to parse than non-primed ones, however again, this appears unrelated to a sluice or parallel prediction. In order to be completely certain about our default prediction tendencies, however, at least insofar as sluicing is concerned, one final study which would ideally be required is a sentence completion (SC) one, similar to the one Yoshida et al. (2013) also conducted. The idea behind such a study would be to investigate how often participants complete the partial utterances provided with a sluice or parallel structure of their own accord. Furthermore, by including pied-piped PPs (which are either primed or not) at the end of such partial utterances, we can gain more information about how such Ps may be used to inform our predictions.

Aside from gaining further evidence for or against a sluice or parallel prediction in this manner, it is also particularly interesting to investigate this behaviour in light of a noisy channel model which makes specific predictions about how we may sometimes overlook a missing P or substitute one where deemed necessary, depending on what structure has the largest prior likelihood of occurring given the context. In other words, this sentence completion task may give us more insight into whether sometimes attaining a higher overall posterior probability may in fact encourage active minor edits to the available material to hand. More specifically, it would be interesting to see whether sometimes participants might be willing to overlook these Ps in favour of a continuation which does not contain them, with such a continuation presumably being more probable than one which *does* contain them. These questions are what the next study attempts to address.

4.5 Sentence Completion Study

4.5.1 Experimental Outline, Aim and Predictions

As mentioned earlier on, in section 4.2, Yoshida et al. (2013) ran a follow-up SC study in which they presented participants (who had not taken part in the original SPR task) with truncated versions of their SPR stimuli. Specifically, these fragments ended just before the reflexive, as shown in (118), repeated below. As the reflexive was missing, the gender mismatch manipulation originally included in their SPR stimuli was nullified.

(118) a. **Condition 1. (Wh-NP)**
Marie's grandfather told some stories at the family reunion, but we couldn't remember which stories about ____.

b. **Condition 2. (Wh-NP)**
Marc's grandmother told some stories at the family reunion, but we couldn't remember which stories about ____.

c. **Condition 3. (Wh-PP)**
Marie's grandfather told some stories at the family reunion, but we couldn't remember with which stories about ____.

d. **Condition 4. (Wh-PP)**
Marc's grandmother told some stories at the family reunion, but we couldn't remember with which stories about ____.

The goal of this fragment completion task was to gain further evidence in favour of a default tendency to predict sluicing wherever possible. Furthermore, they hoped to gain more support for their hypothesis that a pied-piped P, as in the Wh-PP conditions, was enough to block such a sluice projection. Specifically, their hypothesis for this study was that the Wh-NP conditions would be more frequently completed with a sluice compared to the Wh-PP ones.

They found that, on average, participants were indeed much more likely to complete the Wh-NP conditions with a sluice (defined as a ‘single-word completion’ for the purposes of the study) (65.1% of completions), compared to the Wh-PP conditions, for which a sluice would be ungrammatical. However, sluice-style completions were still obtained even for the Wh-PP conditions (15.7%). Although Yoshida et al. take this to simply be experimental ‘noise’, i.e. mistaken readings or chance completions on behalf of the participants, this an intriguing finding which I intend to explore in more depth below.

Given the Greek and simplified English SPR studies we have presented thus far were replication attempts of the original Yoshida et al. study, it stands to reason to also conduct a SC study to compare our results. Aside from this quite basic motivation, however, as also mentioned at the end of the previous experiment’s Discussion section, this SC task is important for two further reasons. Firstly, considering neither the Greek SPR task (4.3) nor the simplified English one (4.4) found any evidence for sluicing or even a parallel structure being projected by the parser, this task represents one final attempt to find any support for such a hypothesis being warranted. Furthermore, given we also found evidence that parsing pied-piped Ps is comparatively harder when they have not been primed, such a SC task can provide us with further interesting information regarding how the parser deals with these pied-piped Ps. Bearing in mind the fact that we are entertaining a noisy channel hypothesis of sluicing whereby prior probabilities play against probable string edits to reach the highest possible posterior probability of each structure, this study can shed further light on whether the prior probability of a sluice or other continuation can lead the parser to ignore or overlook these pied-piped Ps when they do not match this continuation.

With respect to the first goal, gathering evidence for a sluice projection, the hypothesis is straightforward. When presented with a sluice-compatible sentence fragment, participants should complete this fragment significantly more frequently with a sluice than with any other type of continuation. If they have a tendency to project parallel structures, e.g. by simply repeating the information contained in the fragment, then they should complete the utterance significantly more frequently with such a parallel structure instead. This particular iteration of the SC task, however, differs in one key aspect from the original Yoshida et al. one. Specifically, the original study ended each fragment with a *wh*-phrase, repeating the indefinite contained in the potential antecedent (e.g. *. . . some stories. . . , but we couldn’t remember which stories* __). By providing a *wh*-phrase repeating this NP, we believe that the original SC fragments were too restrictive, essentially feeding participants with a ready sluice or the beginnings of a parallel structure. As such, they do not technically represent impartial results on the true frequency of sluice or parallel structure completions. In order to truly support a hypothesis of our default propensity for sluicing, or ellipsis more generally, then we should be able to find evidence of such projections in a minimally appropriate context without such explicitly guiding cues, i.e. in the context of the contrastive conjunction followed by a knowledge verb (e.g. *. . . but we couldn't remember* __). For this reason, we decided to truncate the English SPR stimuli *before* the *wh*-phrase, resulting in the four conditions presented in (4.10).

Table 4.10: SC Study Conditions and Examples

	Condition	Example
1	Wh-NP, P-less Verb	Andrew heard some rumours at the pub, but I do not know for sure ____.
2	Wh-PP, P-less Verb	Andrew heard some rumours at the pub, but I do not know for sure **to** ____.
3	Wh-NP, PP-Verb	Andrew listened to some rumours at the pub, but I do not know for sure ____.
4	Wh-PP, PP-Verb	Andrew listened to some rumours at the pub, but I do not know for sure **to** ____.

Explanatory Note: Bold emphasis included here for presentation purposes only.

The rate of completion of each of these conditions will provide valuable information with respect to any natural prediction tendencies we may have. If, despite the SPR results presented above, the Yoshida et al. (2013) hypothesis is correct and we do have a natural propensity for sluicing, then we would anticipate that in all conditions except the Wh-PP, P-less Verb condition we should have significantly more sluice completions than any other type of completion. If this is true, then we would need to find an alternative explanation for why the SPR results did not show any evidence of reanalysis when such a sluice projection was disambiguated. Similarly, if we have a natural propensity for parallelism instead of ellipsis, then we would anticipate that in all conditions other than the Wh-PP, P-less Verb condition, we would find significantly more parallel structure completions than any other type of completion. Again, if this proves to be true, then we would need to find another explanation as to why the SPR results did not show any slow-down when such a parallel prediction was disambiguated.

With respect to P-pied-piping there are multiple predictions here that need to be carefully considered and teased apart. The first prediction one can make is about how the parser treats such pied-piped Ps with respect to whether their presence or absence are taken into consideration at all to help shape our predictions, as both the SPR studies reported here and the original Yoshida et al. (2013) study take for granted. Specifically, if we seriously take into consideration the presence or absence of these Ps to help shape our predictions, then we would anticipate a significant interaction between the factors of Wh-Type and Verb

Type, such that the Wh-PP, P-less Verb condition has significantly fewer sluice or parallel completions compared to any other condition given it is the only one incompatible with these continuations. Furthermore, even though all three of the remaining conditions are compatible with both a sluice or parallel continuation, if we use the presence of Ps to help guide us, then perhaps we would anticipate the Wh-PP, PP-Verb condition to have significantly *more* sluice or parallel completions compared to the other two.

That having been said, there is also the opinion in the literature, as mentioned before, that pied-piped PPs in English embedded clauses are hard, if not impossible, to parse (Cable & Harris, 2011; Heck, 2008). If, despite what we found in the simplified English SPR study, this is actually true, then we would predict a different set of results. Specifically, neither of the Wh-PP conditions should engender a sluice or parallel continuation, regardless of whether the fragment-final P can be easily incorporated into such a continuation or has been previously primed, i.e. regardless of Verb Type. As such, if we do indeed have a default tendency towards sluicing or parallel structures, then we would anticipate a main effect of Wh-Type, with Wh-NP conditions having significantly more sluice or parallel structure continuations than Wh-PP conditions; we would, furthermore, not anticipate a significant interaction of Wh-Type with the factor of Verb Type, i.e. Verb Type should not have a significant effect on this difference. If, again, we do not have a default sluice/parallel structure tendency, then we would not anticipate either a significant main effect or interaction, with no condition being significantly more prone towards these structures than any other.

In all of the above predictions, we have assumed that the presence or not of a fragment-final P, along with whether that P can be more or less easily incorporated into the 'antecedent' text, will be parsed and taken into consideration in order to inform participants' fragment completion decisions. However, we must also consider the possibility that participants may subconsciously opt to ignore this fragment-final P when its prior probability of occurrence is very low, whilst the prior probability of other types of non-P-pied-piping continuations, such as a sluice or parallel structure, is very high, resulting in a higher posterior probability for a non-P-pied-piping continuation. In other words, we must examine how the prior probability of sluicing or parallelism, but also P-pied-piping, may affect participants' perception of bottom-up material. Specifically, if, on the one hand, we have a tendency towards sluices or parallel structures, i.e. the prior probability of these structures is higher than that of any other structure, whilst at the same time the prior probability of a pied-piped P is very low (because they are infrequent and/or unacceptable), then we may find participants *ignoring* the fragment-final P in the Wh-PP, P-less Verb condition and completing it with a sluice or parallel continuation regardless, given this completion would represent an overall higher posterior probability than one containing a pied-piped P. In other words, we would anticipate a similar, high rate of sluice/parallel completions across all four conditions regardless of Wh-Type, with no main effects or interactions in any direction. Following the same thread of logic, for the Wh-PP, PP-Verb condition in particular, we may also find participants ignoring the fragment-final P, completing the fragment with a parallel structure and then (ungrammatically) *repeating* the P as though P-stranded, e.g. *Andrew listened to some rumours at the pub, but I do not know for sure to which rumours he listened to.* On the other hand, it could be the case that we have no inherent propensity towards either sluicing or parallelism. If this is true, then following the same logic as above, a non-sluice non-P-pied-piping continuation may still represent a higher overall posterior probability than a P-pied-piping continuation, meaning that participants complete all four conditions equally frequently with some other, non-sluice and non-parallel continuation, regardless of the presence or not of a fragment-final P. This would translate to no main effects or interactions of factors in any direction and an overall low sluice and/or parallel structure continuation rate. This examination of how participants may weigh up the posterior probability of a pied-piped PP against that of a string-edited, non-P-pied-piping structure is particularly important, since if either of the above considered situations is true and participants do indeed show a tendency to ignore fragment-final Ps in favour of more probable continuations, then one would need to seriously reconsider the efficacy of using 'pied-piped' PPs to guide participants towards or against certain structural predictions, as all three of the previous SPR studies considered here have done.

4.5.2 Methods

4.5.2.1 Method

This was a web-based study, designed and hosted on Gorilla™(www.gorilla.sc)[17], with participants recruited through Prolific Academic (www.prolific.ac). Participants were given a digital copy of a study information sheet and consented to taking part knowing that the study posed no foreseeable risk to them and that they were able to withdraw at any time without repercussions.

Participants would see a sentence fragment appear on screen and were asked to complete the sentence however they thought appropriate. Null answers were not accepted. Five practice items were presented before the experimental items, and truncated versions of the fillers from the SPR study (see below) were also interleaved among experimental stimuli. Stimuli were presented in a Latin square design using four pseudo-randomised lists, with equal numbers of participants recruited per list (14).

4.5.2.2 Participants

There were 56 participants recruited through Prolific Academic. Subjects were healthy, adult, native speakers of English, aged 18 - 50, with normal or corrected-to-normal vision,

[17]To cross-reference with release notes, data was collected between 1 July 2018 and 15 July 2018.

no hearing, language or speech impairments and no prior history of psychiatric illness. The study took approximately 30 minutes to complete and participants were paid £4 for their time.

4.5.2.3 Stimuli

All 32 experimental items were truncated versions of those presented in the simplified English SPR study (4.4). There were also 64 fillers, which were mostly truncated versions of those used in the SPR task, with the exception of the sluice fillers, which were adapted to simple actives or passives. All fillers used were either truncated simple active, simple passive or relative clauses. Experimental and filler stimuli were all approximately the same length (Exp. Stimuli Mean: 13.97; SD: .2; Fillers Mean: 13.63; SD: 2.3).

4.5.3 Results

4.5.3.1 Completion Categorisation

As mentioned above, participants were required to respond with at least one word, with null answers not being accepted. We then classified these responses into 4 categories: ellipsis; clefts; parallel continuation; other; as defined below:

- **Ellipsis**
  - **Sluicing:** Reducing an embedded *wh*-question to its initial *wh*-element, with the remnant matching the antecedent correlate and the antecedent verb being easily inferred.

    (144) Edward played some songs on the piano, but I do not recall with certainty **which songs**.

  - **Sprouting:** Reducing an embedded *wh*-question to its initial *wh*-element, with the remnant not corresponding to a correlate in the antecedent, and the antecedent verb being easily inferred.

    (145) Edward played some songs on the piano, but I do not recall with certainty **why/how/for how long**.

- **Cleft**
  Repeating the antecedent correlate NP, or referring to it in some fashion (*wh*-element, pronoun etc.), followed by a copula.

  (146) Edward played some songs on the piano, but I do not recall with certainty **what those songs were/what songs they were/what these were**.

- **Parallel**
  Repeating both the antecedent verb and correlate NP (overtly or as interrogative pronoun, simple pronoun does not suffice).

  (147) Edward played some songs on the piano, but I do not recall with certainty **which songs he played/what he played**.

- **Other**
  Any other type of completion, such as embedded questions, relative clauses, other modifying phrases etc.

  (148) Edward played some songs on the piano, but I do not recall with certainty **who had asked for them/if the songs were sad/what the others thought of them/at what time that was/etc.**

With respect to the Wh-PP conditions, i.e. those ending in a P, there were two further classifications which could be made: a) either the continuation acknowledged the fragment-final P and grammatically incorporated it, e.g. by using the P to create a sluice or parallel continuation in the Wh-PP, PP-Verb condition (149), or by using the P to create some form of modifier for either Wh-PP condition (150); or b) the continuation ignored the P, instead continuing with an ungrammatical completion, e.g. making an otherwise well-formed elliptical or parallel continuation for the Wh-PP, P-less condition (151) etc.

(149) Anna listened to some songs on the radio, but I do not recall with certainty to **which songs/which songs she listened/what extent she had heard them before/etc.**

(150) Edward played some songs on the piano, but I do not recall with certainty for **what reason he played them/how long he played/etc.**

(151) * Edward played some songs on the piano, but I do not recall with certainty for **which songs/what songs he played/why/etc.**

Based on this classification system, table 4.11 and figure 4.8 show the percentage of each type of continuation for each condition, with more detailed analyses following.

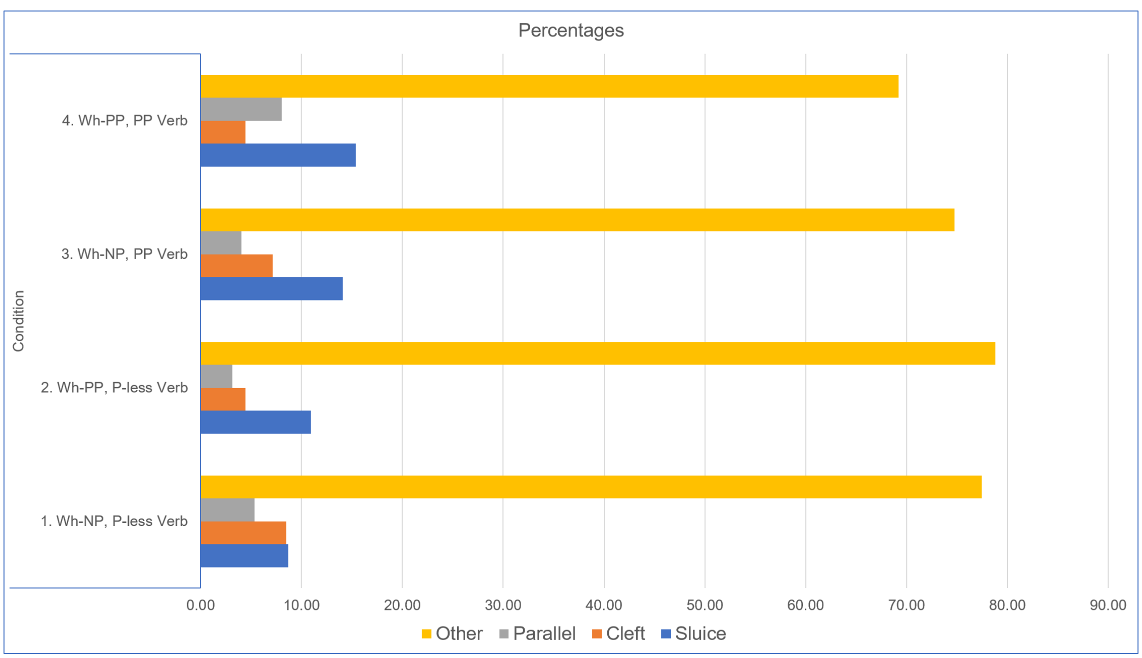
4.5.3.2 Analysis

Based on this result classification system, we sought to answer two main questions: a) whether we project sluicing (or parallel or cleft) continuations more than any other type of continuation when the context allows for it (i.e. in sluice-compatible conditions); and b)

Table 4.11: SC Results; Continuation Percentages by Condition

Response	1. Wh-NP, P-less Verb	2. Wh-PP, P-less Verb	3. Wh-NP, PP-Verb	4. Wh-PP, PP-Verb
	Condition			
Sluice	8.71	10.94	14.06	15.4
Cleft	8.48	4.46	7.14	4.46
Parallel	5.36	3.13	4.02	8.04
Other	77.46	78.79	74.78	69.2

Figure 4.8: Sentence Completion Results; Percentages by Condition



whether we use PPs to inform these continuation decisions, i.e. whether a fragment-final P is enough to discourage such continuations.

To answer both these questions, a Chi-Square analysis was run in order to establish whether there was a significant correlation between continuation type and condition. Results indeed showed a significant correlation ($\chi^2(9) : 33.91; p < 9.27e{-}05$). This result was broken down further using post-hoc Bonferroni-corrected pairwise analyses, shown in detail in Table 4.12 (with boxes colour-coded by continuation type and significant results highlighted in red lettering).

These analyses found that by far the most frequent type of continuation across all conditions was 'Other' (all p's$< 2e - 16$), and not elliptical, parallel or cleft continuations.

Among these three remaining continuations, analyses showed that elliptical ones were more frequent than parallel and cleft ones in *some* conditions, but not the expected ones. Overall, there was no significant difference between cleft and parallel continuations for any condition (all p's$> .105, n.s.$)

Specifically, breaking continuations down by Condition, for the sluice compatible Wh-NP, P-less Verb condition, there was no significant difference between any of the three continuations (all p's$> .071, n.s.$). For the sluice *non*-compatible Wh-PP, P-less condition, on the other hand, there were actually significantly *more* sluices than parallel continuations ($p < .001$) and marginally more than clefts ($p < .057$), against expectations. Numerically, there were actually more sluice continuations for this condition than the sluice compatible Wh-NP, P-less Verb condition, although this difference was not significant (see confirmatory logistic regression analysis below). For the sluice compatible Wh-NP, PP-Verb condition, there were significantly more elliptical than parallel continuations ($p < 2.80e - 06$), but not clefts ($p > .18, n.s.$). Finally, for the sluice compatible Wh-PP, PP-Verb condition, there were significantly more elliptical than both parallel ($p < 2.5e - 05$) and cleft continuations ($p < 2.5e - 05$).

These differences were also supported by logistic regression analyses, conducted to follow the Yoshida et al. (2013) analyses. Specifically, three separate mixed effects logistic regression models were fitted to the data with sluice probability (1,0); cleft probability (1,0) and parallel probability (1,0) as DVs and Verb Type (P-less Verb vs. PP-Verb) and Wh-Type (Wh-NP vs. Wh-PP) as fixed effects IVs in each, with subjects and items as random effects (random slopes and intercepts assumed). For the sluice continuations, results unexpectedly showed a highly significant main effect of Verb Type ($z < 3.36; p < .001$), with PP-Verbs eliciting significantly more elliptical continuations (132) than P-less Verbs (88). Interestingly, this effect was not modulated by an interaction with the factor of Wh-Type ($z < 1.85; p > .064, n.s.$), nor was there a main effect of Wh-Type ($z < 1.10; p > .27, n.s.$), with no difference between conditions ending in a P and those not ending in a P with respect to the number of sluice completions, regardless of their compatibility or not with such a sluice completion. This indicates that a fragment-final P had no effect on sluicing continuation availability. For cleft continuations, results showed a significant main effect of Wh-Type ($z > 2.99; p < .003$), with Wh-NP conditions being more frequently completed with a cleft (70) compared to Wh-PP conditions (40), with no main effect of Verb Type ($z < 0.51; p > .61, n.s.$) nor a significant interaction of Wh-Type and Verb Type in any direction ($z < .552; p > .581, n.s.$). For parallel continuations, results showed no main effects (both $z's < 1.63; p > .104, n.s.$), however there was a significant interaction between the two factors ($z = 3.10; p < .001$), with the difference between Wh-NP and Wh-PP conditions being affected by Verb Type such that in the PP-Verb conditions, the Wh-PP condition had significantly more cleft continuations than the Wh-NP one, whereas in the P-less Verb conditions, the Wh-PP condition had significantly fewer cleft continuations than the Wh-NP one. For both clefts and parallel continuations, therefore, it does look as though fragment-final Ps were taken into consideration to inform participants' continuation choice, even though this was apparently not the case with sluice continuations.

Finally, we examined how frequently participants accommodated for fragment-final Ps, incorporating them into a grammatical continuation, and whether this frequency was affected by having been primed by the same P in the antecedent (i.e. in PP-Verb conditions). Isolating the Wh-PP conditions, a mixed effects logistic regression model was fitted to the data with P accommodation probability (1, 0) as the DV and Verb Type ($\pm$PP) as the sole fixed effect, with subjects and items as random effects (random slopes and intercepts assumed). Results showed no effect of Verb Type ($z = .74; p > .46, n.s.$), indicating that being primed by the P in the antecedent had no effect on how easily participants integrated it into their continuation across all types of continuation. Numerically, there was a difference between the overall frequency of grammatical accommodation (Wh-PP, P-less Condition: 271; Wh-PP, PP-Verb Condition: 261; Overall: 531) and non-accommodation (Wh-PP, P-less Condition: 165; Wh-PP, PP-Verb Condition: 175; Overall:340), however this difference was not significant ($p > .51, n.s.$).

4.5.4 Discussion

Overall, the results of the SC task showed that participants did not have a clear preference for sluicing, clefts or parallel structures, with a significantly larger number of continuations produced being structures 'other' than these. Furthermore, in the one condition where all the SPR studies assumed we would automatically predict sluicing, i.e. the Wh-NP, P-less Verb condition, there was no significant difference between elliptical, parallel structure or cleft continuations, indicating that any of these is actually a plausible prediction. As such, this study adds further evidence against the hypothesis that we have an in-built tendency for sluicing, and ellipsis more generally, when the context allows for it. These results also indicate that the original Yoshida et al. (2013) SC task findings were likely driven by the

Table 4.12: SC Chi-Square Results

	Parallel (Wh-PP; P-less Verb)	Parallel (Wh-NP; PP-Verb)	Cleft (Wh-PP; P-less Verb)	Cleft (Wh-PP; PP-Verb)	Parallel (Wh-NP; P-less Verb)	Cleft (Wh-NP; PP-Verb)	Parallel (Wh-PP; PP-Verb)	Cleft (Wh-NP; P-less Verb	Ellipsis (Wh-NP; P-less Verb)	Ellipsis (Wh-PP; P-less Verb)	Ellipsis (Wh-NP; PP-Verb)	Ellipsis (Wh-PP; PP-Verb)	Other (Wh-PP; PP-Verb)	Other (Wh-NP; PP-Verb)	Other (Wh-NP; P-less Verb)
Parallel (Wh-NP; PP-Verb)	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Cleft (Wh-PP; P-less Verb)	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-
Cleft (Wh-PP; PP-Verb)	1	1	1	-	-	-	-	-	-	-	-	-	-	-	-
Parallel (Wh-NP; P-less Verb)	1	1	1	1	-	-	-	-	-	-	-	-	-	-	-
Cleft (Wh-NP; PP-Verb)	0.95465	1	1	1	1	-	-	-	-	-	-	-	-	-	-
Parallel (Wh-PP; PP-Verb)	0.22354	1	1	1	1	1	-	-	-	-	-	-	-	-	-
Cleft (Wh-NP; P-less Verb	0.10489	0.90316	1	1	1	1	1	-	-	-	-	-	-	-	-
Ellipsis (Wh-NP; P-less Verb)	0.07137	0.64926	1	1	1	1	1	1	-	-	-	-	-	-	-
Ellipsis (Wh-PP; P-less Verb)	0.00124	0.01828	0.0577	0.0577	0.41198	1	1	1	1	-	-	-	-	-	-
Ellipsis (Wh-NP; PP-Verb)	2.80E-06	6.90E-05	0.00028	0.00028	0.00348	0.17641	0.79867	1	1	1	-	-	-	-	-
Ellipsis (Wh-PP; PP-Verb)	1.90E-07	5.50E-06	2.50E-05	2.50E-05	0.00037	0.02781	0.15358	0.32731	0.46709	1	1	-	-	-	-
Other (Wh-PP; PP-Verb)	2.00E-16	2.00E-16	2.00E-16	2.00E-16	2.00E-16	2.00E-16	2.00E-16	2.00E-16	2.00E-16	2.00E-16	2.00E-16	2.00E-16	-	-	-
Other (Wh-NP; PP-Verb)	2.00E-16	2.00E-16	2.00E-16	2.00E-16	2.00E-16	2.00E-16	2.00E-16	2.00E-16	2.00E-16	2.00E-16	2.00E-16	2.00E-16	1	-	-
Other (Wh-NP; P-less Verb)	2.00E-16	2.00E-16	2.00E-16	2.00E-16	2.00E-16	2.00E-16	2.00E-16	2.00E-16	2.00E-16	2.00E-16	2.00E-16	2.00E-16	1	1	-
Other (Wh-PP; P-less Verb)	2.00E-16	2.00E-16	2.00E-16	2.00E-16	2.00E-16	2.00E-16	2.00E-16	2.00E-16	2.00E-16	2.00E-16	2.00E-16	2.00E-16	1	1	1

wh-phrase providing an explicit sluice cue.

Besides this interesting finding, however, another very important result was that the presence or not of a fragment-final P did not appear to affect participants' sluice continuation tendencies at all. That is to say, participants were equally likely to complete the utterance with a sluice regardless of whether a fragment-final P would make it ungrammatical or not. What appeared bizarrely to significantly affect the likelihood of a sluice continuation, instead, was the factor of Verb Type, with PP-Verbs being significantly more likely to encourage a sluice continuation compared to P-less verbs, regardless of Wh-Type. This is quite a perplexing finding which would require further study to replicate and investigate more fully. However, the lack of interaction with or main effect of Wh-Type here, with the one sluice non-compatible condition (Wh-PP, PP-Verb Verb) not actually having significantly fewer sluice continuations than any other condition, does show that using a pied-piped P does not necessarily discourage a sluice projection, a basic assumption in all the SPR studies discussed here. We must, therefore, conclude that the original Yoshida et al. manipulation, used also in both the Greek and English follow-up SPR studies, although undoubtedly theoretically sound, quite likely did not affect participants in the desired way, i.e. it did not serve to foil a sluice projection.

On the other hand, Wh-Type *did* affect the likelihood of both parallel and cleft structure continuations, such that a) Wh-NP conditions were significantly more likely to be continued with a cleft compared to Wh-PP conditions, with a main effect of Wh-Type; b) Wh-PP, PP-Verb conditions were significantly more likely to be continued with a parallel structure compared to Wh-PP, P-less Verb conditions, exhibited as an interaction of Wh-Type*Verb Type. In other words, in these two situations participants appear to be using the fragment-final P to inform their continuation as we would expect them to, even though they did not do so with sluices. As such, the manipulation of Wh-Type in the SPR studies, with P-pied-piping vs. bare argument integration, appears to have been much more likely to discourage a parallel or cleft continuation compared to a sluice one as was the original intention.

Regardless of the insight these results afford us with respect to the SPR data, however, this difference between clefts, parallel structures and sluices in terms of whether they take into consideration or ignore this fragment-final P is also very interesting from a noisy channel perspective, as it lends evidence to the idea of the parser being willing to ignore Ps when it comes to sluicing, but less so when it comes to clefts or other overt continuations. This finding is intriguing since it may contribute to the idea that sluicing in particular, or ellipsis more generally, is processed differently from other, overt structures, with the parser being willing to accommodate or overlook what otherwise constitute ungrammatical structures, at least insofar as the presence or absence of Ps is concerned. Whether this accommodating nature towards sluicing is related to the idea of ellipsis being treated as a noisy channel in communication or whether it is Ps, in particular, which are for some reason oddly parsed remains to be confirmed by further studies.

Finally, overall, in terms of accommodation probability, i.e. how likely participants were to generally somehow grammatically incorporate the fragment-final P in their continuation, there appeared to be no significant advantage for PP-Verbs over P-less Verbs. In other words, whether or not participants had already seen the P in the 'antecedent' clause or not did not affect how likely they were to incorporate the fragment-final P.

4.6 Overall Discussion

At the end of this chapter, we have collected a very intriguing set of results, allowing us to draw some interesting conclusions. Our motivation going in to this chapter was to supplement the off-line acceptability data collected from two overtly non-P-stranding languages, namely Greek and German, with on-line data regarding the parser's behaviour with respect to P-stranding under sluicing. Specifically, the explanation provided in Chapter 2 for these off-line data is based on the notion that the parser treats sluicing as a noisy channel, calculating the posterior probability of the intended message based on the prior likelihood of this message along with the probability that this intended message may have been corrupted into the perceived one. This probability of 'corruption' is linked firstly to how noisy the context the message appears in is; by treating sluicing as a noisy channel, this opens up the possibility for the perceived sluice remnant to be a slightly corrupted version of the producer's intended message, compared to, e.g., overtly expressed material where no noise is considered present. Secondly, this probability of corruption is tied to the number of string edits required to match the perceived message with the hypothetical intended one. The lower the number of string edits required, the higher the likelihood of corruption, with 0 string edits (i.e. no corruption) being the most likely. The higher the number of string edits, on the other hand, the lower the likelihood of corruption and the larger the likelihood that the parser will have to consider a different hypothetical intended message which requires fewer string edits to reach. Together, the prior likelihood of the intended message and the likelihood of corruption of this intended message into the perceived one form the posterior likelihood of the intended message, with a higher posterior likelihood making an intended message more likely to be inferred by the parser based on the perceived message compared to one with a lower posterior likelihood.

However, even in situations where the context is noisy and therefore corruption is considered likely, such as we hypothesize is the case for sluicing, each string edit required to reach the intended message from a corrupted version of this message is further hypothesized to carry with it an associated processing cost. Therefore, an intended message which is further away from the perceived one in terms of number of string edits should theoretically be more costly to parse compared to one which is closer or even identical to the perceived one. The phenomenon of P-stranding under sluicing in non-P-stranding languages is considered to be one such instance of corruption of the intended message, where the parser treats P-stranding as simply an instance of P-pied-piping where the P has been deleted due to noise. In order to reach the intended P-pied-piping message from the perceived P-stranding one, 1 string edit (i.e. adding a P) is required. This should make parsing a perceived message with P-stranding under sluicing more costly compared to one with P-pied-piping under sluicing, despite both having the same intended message. This difference in processing costs is then translated to different acceptability ratings, with larger costs leading to lower acceptability, and lower costs to higher acceptability. This explains the consistent pattern we found across both Greek and German, where a) P-stranding under sluicing was significantly more acceptable than P-stranding overtly (thanks to sluicing making corruption more likely); and b) P-stranding under sluicing was significantly less acceptable than P-pied-piping under sluicing (thanks to the former requiring one string edit more than the latter in order to be properly parsed, thereby entailing a higher processing cost).

The aim of this chapter, therefore, at the outset, was to gather on-line evidence in favour of such a hypothesis by identifying the processing costs associated with P-stranding vs. P-pied-piping under sluicing in overtly non-P-stranding languages and establishing whether the former is indeed harder to parse in real time compared to the latter. In order to do so, a particular SPR study by Yoshida et al. (2013), originally run in English, was identified as a suitable candidate to demonstrate these processing costs, measured as differences in reading times. This study claimed that we have an 'innate' tendency to project sluicing whenever presented with the opportunity, with this claim being based on our behaviour towards a reflexive pronoun supposedly contained within such a projected sluice and constrained by the rules of binding. Specifically, the study showed that whenever a sluice was allowed to be predicted, if the gender of this reflexive made it not fit this prediction, then participants would react, exhibiting a GMME; on the other hand, whenever a sluice was blocked, the gender of the reflexive did not appear to matter at all to participants. If this predisposition to generate sluice predictions is true, then it may be possible for us to use this study design in, e.g., Greek and manipulate it in order to generate such sluicing predictions and establish how the parser behaves when it encounters a P-stranding remnant when it should anticipate P-pied-piping, vs. when it encounters P-pied-piping.

There were two possible theoretical issues with the original Yoshida et al. (2013) design, however. The first issue was that, based on the study design and due to English not having overt case-marking, when a sluice prediction was possible, it was equally plausible that participants may be predicting parallel structure continuations or clefts instead of or as well as sluices, with no way of distinguishing between the three. The second issue concerns the study's use of reflexives contained within representational NPs (RNPs), such as *picture of myself*. Specifically, these RNP-contained reflexives have been argued in the theoretical literature to be *exempt* from regular binding restrictions. That having been said, if this was indeed true, then it was unclear why the original study only found a GMME when a sluice was predicted, i.e. when these binding restrictions were present, as opposed to when it was not. As such, leaving aside this particular theoretical objection, we considered this a reliable design to use, intending to address the first issue in due course.

Our attempt to replicate this experiment in Greek, however, failed. Specifically, it appeared as though the reflexive in Greek generated a GMME regardless of sluice predictability. This result was concerning and raised the question of whether the original Yoshida et al. (2013) study result had, in fact, been driven by other factors. Specifically, we hypothesized, firstly, that the RNP-contained reflexive is not, in fact, constrained by the rules of binding, as had been argued in the literature, leading it to generate a GMME irrespectively of sluice predictability; secondly, we identified a local attachment ambiguity in the original design which may have garden-pathed or confused the participant in the sluice non-compatible conditions and prevented the GMME from being computed in these, leading to the observed results. Importantly, this local attachment ambiguity is absent in Greek, hence the difference in results between the two studies.

Despite this issue, this result did not mean that the original study claim was false. Indeed, it was still possible that we may project sluicing by default whenever given the opportunity, we simply had to measure this tendency differently, in a way that removed these confounds. If this tendency does exist, then we may still be able to eventually use a modified version of the Yoshida et al. (2013) study for our purposes, i.e. to measure the processing costs of P-stranding vs. P-pied-piping under sluicing in Greek. For this reason, we created a modified version of the original study, excluding this reflexive and removing the local attachment ambiguity, and ran it in English in order to gather evidence in favour of such a sluice prediction tendency. Despite this modification, however, this new study also returned a null result, finding no evidence for an natural preference for sluicing projections.

Given all three of these study designs, from the original Yoshida et al. one, to the Greek and eventually the modified English one, were based on the presence or not of a pied-piped P to either allow or block a sluice prediction, we also took the opportunity in this last study to investigate one more claim made in the theoretical literature regarding English P-pied-piping. Specifically, it has been claimed that P-pied-piping in embedded clauses is much less frequent in English and harder to parse compared to P-stranding. By manipulating the ease of integrating such a pied-piped P in this last English study, we took the opportunity to observe how participants processed a) a *wh*-PP when it could be grammatically integrated into a sluice prediction, i.e. how regular P-pied-piping is parsed; vs. b) a *wh*-PP when it could not be grammatically integrated into a sluice prediction, i.e. how unexpected P-pied-piping is parsed. If the theoretical literature re the unacceptability of P-pied-piping in English is correct, then we would not expect a significant difference between these two.

Instead, what we found was that integrating an expected P was significantly easier than an unexpected one. This result, however, did not appear to be driven by sluice predictability, as there was no significant interaction observed at a later point in the sentence where such a sluice projection would be disambiguated. Instead, we concluded it was driven rather by a priming effect, with an already encountered P being easier to parse the next time round compared to a P encountered for the first time. This last result, in particular, also contributed to our overall conclusion that participants did not appear to be actively predicting sluices.

To supplement this last study and gather one final piece of information in favour of or against default sluicing projection tendencies as well as regarding how we treat pied-piped Ps, we ran a SC study using truncated versions of the last English SPR study's stimuli, similarly to a follow-up SC study also run by Yoshida et al. (2013). Specifically, we hypothesized that if the parser has a tendency for sluicing, then we would anticipate significantly more sluices than other types of continuation whenever sluicing is allowed. This study could also help shed light on one of our original objections to the SPR study's design, namely whether participants were indeed projecting sluice continuations where claimed, or whether perhaps they were entertaining parallel structure or cleft continuations instead. If we found sluices to be a significantly more frequent continuation where claimed in this SC study, then this would aid in dispelling this original objection. Finally, with respect to the presence or not of Ps indeed serving to allow or block a sluice - or parallel structure or cleft continuation - in the SPR studies as predicted, we would anticipate this SC study to find a significant difference in the number of sluice completions when the sentence fragment ended in such a sluice-incompatible P vs. when not.

In harmony with our other study results, the results of the SC study showed that, overall, we do not appear to have a tendency to predict sluicing, with other types of continuations (excluding parallel structures and clefts) being significantly more frequent than sluices. Furthermore, with respect to the probability of sluice vs. parallel structure vs. cleft projections in the SPR studies, the SC study found that in the condition where we supposedly project sluicing (namely the Wh-NP, P-less Verb condition) there was no significant difference between the three. This result indicates that, at least as far as English is concerned, we cannot technically distinguish between any of these structures as a plausible projection where the SPR studies originally argued only for a sluice projection. This finding, therefore, provides support in favour of our first objection to the original SPR study's claims.

Finally, the presence or not of a fragment-final sluice-blocking P when the antecedent did not contain the same P did not appear to affect these sluice continuations whatsoever, with sluices being equally frequent either way. On the other hand, it *did* appear to have an effect on the frequency of parallel or cleft continuations in the expected direction, i.e. with fragment-final Ps not contained in the antecedent discouraging both types of continuation.

This last, intriguing result serves, firstly, to show that the original SPR study manipulations based on P-pied-piping likely did not have the desired effect of blocking a sluice prediction in some conditions vs. others. Secondly, however, it also provides further evidence in favour of Ps being apparently processed differently in the context of sluicing compared to overt contexts, a finding in-line with our hypothesis that sluicing may be treated as a noisy channel.

In conclusion, although we were not able to isolate and identify the processing costs associated with P-pied-piping vs. P-stranding under sluicing in overtly non-P-stranding languages, as had been the original goal of this chapter, the above presented studies have illuminated a problem in the existing sentence processing literature with respect to sluicing and helped unravel its cause. As such, I believe their results are an important contribution to this literature, providing evidence that we do not, in fact, appear to have an 'innate' tendency for sluicing, or ellipsis more generally, as had been claimed previously, at least not with the existing datasets to hand.

Finally, given the time limitations of a PhD, these experimental setbacks meant that we were not able to design an alternative paradigm to identify the online effects of P-stranding vs. P-pied-piping in Greek or German. That having been said, however, we did gather valuable information against relying on pied-piped Ps to discourage ellipsis in our designs moving forward. Instead, an alternative way to gather evidence for a noisy channel hypothesis of sluicing in the future might be to move in a different direction and examine how manipulating the amount of contextual noise present affects how easily participants are willing to accommodate for P-less arguments in overtly non-P-stranding languages. For instance, if sluicing is indeed noisier than fully overtly expressed structures, then we might be able to find evidence for this by comparing the acceptability and/or processing costs associated with P-deletion overtly vs. under sluicing. This idea among others will be examined in more detail in the next chapter, when promising future directions for this research are discussed.

Chapter 5

Conclusions and Future Directions

5.1 Summary

In Chapter 1, we introduced the topic of sluicing, a form of linguistic ellipsis whereby an interrogative clause is reduced to its initial *wh*-element, the sluice remnant. Sluicing is a particularly intriguing linguistic phenomenon, thanks to the features of this remnant appearing to simultaneously support two opposing types of accounts of sluicing; one involving some form of silent structure at the *e*-site, with the remnant arriving at its surface position through regular *wh*-movement; and the other involving no such structure nor movement, with the remnant simply being base-generated in its surface position. More specifically, since sluicing was first introduced to the linguistic community by Ross (1969), this *wh*-remnant has been shown to bear the same case we would expect had the content within the site of ellipsis been overtly expressed with a clause identical to the antecedent (apart from the *wh*-phrase). Later on, Merchant (2001) also argued that if a sluice's correlate was a full PP, then the sluice remnant could only appear without this P if the language in question also allowed overt *wh*-movement to strand Ps. Together, these two form identity generalisations suggest that at least at some point in the derivation there existed a full *wh*-question at the site of ellipsis which was simply deleted prior to pronunciation, making sluicing conform to the rules of overt syntax. At the same time, however, ever since this first introduction to sluicing in 1969, it has also been accepted that the sluice remnant appears to defy normal *wh*-movement rules by systematically escaping what would otherwise constitute syntactic islands had this supposed structure within the *e*-site been overt. These two sets of features have created a variety of approaches to sluicing.

Among the many different theories of sluicing and the different ways of classifying them, Chapter 1 focused on the three most influential theoretical accounts, along with one sentence processing one, splitting them broadly into two groups: movement and non-movement accounts. On the one hand, movement accounts consider the sluice remnant to have once been part of a full interrogative CP, whose IP has been deleted prior to PF. As such, it is subject to regular case-marking and must adhere to the rules of *wh*-movement. Such accounts include Ross's original PF-Deletion Hypothesis, along with its revival via Merchant (2001). The original approach advocated a full interrogative CP, subject to the regular rules of overt syntax, being generated at the site of ellipsis under a strict syntactic identity condition with material preceding it, i.e. the sluice's antecedent. At some point prior to reaching PF, this CP's IP is deleted thanks to this identity with the antecedent. Merchant's (2001) version of PF-Deletion is similar, except for the identity condition required of the sluice with its antecedent. Specifically, instead of a strictly syntactic identity condition, Merchant (2001) argued that the sluice must exist in a relationship of mutual semantic entailment with some part of its antecedent, i.e. that one must be semantically inferable via the other. This two-way entailment successfully captured cases that syntactic identity alone could not, such as, for instance, 'vehicle change', where R-expressions can license the deletion of pronouns. What these movement accounts have in common is that by advocating for the remnant being the result of regular *wh*-movement, the two form identity generalisations naturally follow. An additional, key amendment to Merchant's PF-Deletion Theory by Abels (2016) in the form of the Fit Condition, further accounts for those restricted instances where remnant and correlate need not necessarily match in case, so long as the remnant can slot into the correlate's place in the antecedent resulting in the same meaning without further alterations being necessary.

With respect to the phenomenon of island amelioration, there are broadly two approaches within this movement camp. One approach is that only overtly pronounced island-forming nodes result in unacceptability, therefore deletion of these nodes at PF results in a grammatical utterance. That having been said, under this approach it becomes difficult to explain why other forms of ellipsis containing such island-forming nodes, such as VP-ellipsis, stripping, sprouting, or even multiple sluicing, do not appear to allow island amelioration, with further stipulations being necessary (see Abels (2017); Merchant (2001); Vicente (2018) for more details). The other approach considers what we find under sluicing to not be 'true' instances of island amelioration, but rather island *evasion* in the sense that the *e*-site itself need not necessarily contain an island per se (Abels, 2011; M. Barros et al., 2014). Specifically, this approach stipulates that if a suitable, semantically equivalent paraphrase of the antecedent can be found, which however does not involve an island, then it is perfectly acceptable to suppose that the *e*-site contains this paraphrase instead of a syntactically identical copy of the antecedent. In this way, by assuming that island ameliorating sluices in fact have shorter, non-island-containing sources, the issue of island extraction is circumnavigated altogether. This approach, however, also has its drawbacks in that it can over-generate, predicting island amelioration in situations where it does not appear to be allowed.

Moving on to the non-movement accounts, these take the opposite view that there is no syntactic structure within the *e*-site prior to pronunciation, arguing instead that the remnant is simply base-generated in [Spec, CP]. Chung et al.'s (1995) approach of LF-Copying assumes that the meaning of the sluice is derived via copying of the antecedent IP into the *e*-site after the sluice's pronunciation at PF, i.e. at the level of LF. The pointer account, on the other hand, a sentence processing approach, also stipulates no structure at the *e*-site, but this time without material being copied over to the *e*-site at any point. Instead, the site of ellipsis is argued contain a pointer, a silent processing mechanism which guides the parser back to the antecedent in order to obtain the sluice's meaning. By stipulating no syntactic structure at the *e*-site, both these approaches naturally capture island amelioration under sluicing, however, at face value, they fall short in capturing form-identity effects. In order to capture the case-matching generalisation both non-movement approaches posit an additional strict case-matching requirement between remnant and correlate for sluicing to be licensed. However, this case-matching requirement, in turn, is too restrictive, failing to capture situations where we have case-mismatching under sluicing. Finally, neither account in its original form can account for the PSG.

This leaves us with an assortment of sluicing theories, each with its own strengths and weaknesses, but none of which are perfect. The account to most closely follow the veritable pastiche of features that characterise sluicing, however, is arguably Merchant's (2001) PF-Deletion Hypothesis, together with the Fit Amendment. One of the most important contributions of this approach is the P-stranding Generalisation or PSG, the focus of this thesis. However, despite most languages being accepted to follow this generalisation, over the course of almost two decades there have been accounts of more and more languages appearing to defy it, by allowing P-stranding under sluicing whilst disallowing it overtly. At first glance, these accounts present a serious problem for the PSG. That having been said, some of these exceptions, predominantly in non-case-marking languages, have been maintained to not be true counter-arguments to a PF-Deletion theory, with proponents arguing that they can simply derive of a non-P-stranding source, an approach similar to the island evasion one mentioned above. These alternative sources are predominantly copulas or cleft-like structures (Rodrigues et al., 2009), or resumption (Al Bukhari, 2016; AlShaalan and Abels (2019)). As such, it is asserted that, by not deriving of a P-stranding source, they do not represent true counter-examples to the PSG. However, there are also a number of other languages which cannot be explained via such alternative sources, as they lack the potential for them, largely due to case-marking of the remnant or simply a lack of resumptive pronouns. It is when examining the explanations for each of these languages, in particular, that we observe the cross-linguistic landscape of sluicing descending slowly into theoretical chaos. The first issue apparent is that for each of these languages, a different explanation has been proposed, one, however, that does not successfully extend to all other such languages. As such, we have explanations based on PPs being PF-islands for some, but not other languages; P-stranding being ameliorated via a PF-interface repair process based on [+wh] feature percolation; and a pointer mechanism working only for some, but not other languages, a.o. This means that despite the existence of multiple apparent exceptions to the PSG, when one surveys this literature it becomes abundantly clear that there is no cohesive explanation uniting the behaviour of all these PSG-defiant languages and relating it to the behaviour of PSG-compliant ones. Furthermore, an additional important problem which arises from this literature review is that, with the exception of AlShaalan and Abels (2019) for Saudi Arabic, almost no adequately powered and controlled experimental evidence has been garnered to accurately depict this phenomenon, leaving us in the dark as to how acceptable P-stranding truly is under sluicing and how it relates to a) P-stranding overtly, and b) P-pied-piping under sluicing and overtly. At the end of chapter 1, we are thus left with a tangled and confusing state of affairs vis-á-vis the status of the PSG cross-linguistically.

These two very important issues, namely the lack of experimental evidence depicting the behaviour of P-stranding under sluicing, as well as the lack of a cohesive cross-linguistically applicable explanation of this phenomenon form the basis of this thesis and are tackled in Chapters 2 and 3, respectively.

Specifically, in Chapter 2, novel, powerful and well-controlled experimental datasets are provided depicting the acceptability of P-stranding under sluicing in Greek and German. These two overtly non-P-stranding languages have long been held as strong examples of compliance with the PSG in the theoretical literature. In spite of this, over the course of four web-based acceptability judgement tasks in Greek and German, I provide strong evidence that P-stranding is significantly more acceptable in these languages under sluicing than it is overtly, although it is still significantly less acceptable than P-pied-piping, either overtly or under sluicing. Furthermore, this is true despite a cleft or resumptive alternative source not being available for either of these languages. These experiments also provide the first acceptability data in the literature on the behaviour of case-mismatching and island amelioration under sluicing.

In each of these acceptability judgement tasks, the basic factors of P-stranding (P-pied-piping vs. P-stranding) and Sluicing (Sluicing vs. Overt Continuation of Sluice) are crossed. In the first experiment, the factor of remnant Case-matching (Case-matching of remnant and correlate vs. Case-mismatching) is also added to the design. The results of this study showed significant main effects and two-way and three-way interactions of all factors, with sluicing being overall more acceptable than overt continuations; P-stranding being overall less acceptable than P-pied-piping; case-mismatching being overall less acceptable than case-matching; both case-mismatching and P-stranding becoming significantly more acceptable under sluicing compared to overtly; and case-mismatching receiving its highest rating when combined with P-stranding under sluicing and its lowest when combined with P-stranding overtly. The important take-away from this experiment is that for the first time in the existing literature this is a powerful dataset indicating that P-stranding is indeed significantly better under sluicing compared to overtly, whilst still being significantly less acceptable than P-pied-piping. Furthermore, the design of the experiment was such that it provided conclusive evidence against this P-stranding amelioration under sluicing being due to a) the selectional properties of the antecedent verb making a bare remnant simply an acceptable alternative to a full PP; b) an underlying cleft source; or c) a temporary grammatical illusion, where the bare remnant is temporarily mistaken for the immediately preceding verb's argument. It was also argued that an alternative source featuring resumption is impossible, as is treating the P-less remnant as a case of P-dropping where the bare remnant semantically represents the full PP in a form of phrasal lexicalisation.

As mentioned above, one of the existing explanations for PSG-defiance in other languages is that the PP may be a PF-island ameliorated under sluicing. Leaving aside the issues of why this would be true for some, but not other overtly non-P-stranding languages, as well as why P-stranding is not thus ameliorated across all forms of ellipsis, the second and third experiments in Chapter 2 explore this potential explanation. To begin with, even though it has been widely accepted that sluicing, for the largest part, ameliorates (or evades) islands, there has not yet been any experimental data to show exactly what this amelioration pattern looks like. Specifically, are islands entirely ameliorated under sluicing, i.e. are they as acceptable as a structure containing no island to begin with? Or are they perhaps slightly less acceptable, similarly to how we found P-stranding to be less acceptable than P-pied-piping under sluicing, despite being much more acceptable than overtly? If the latter case is true, then one can see how an argument could perhaps be made for PPs in Greek being simply another island ameliorated under sluicing.

In order to examine this possibility, in Experiment 2 the factors of Sluicing and P-stranding are now crossed in the context of islands. In other words, all items involved extraction out of either a relative clause or a subject island. By comparing the results of this experiment to those of the first one, this would provide us with an idea of a) what island extraction looks like under sluicing compared to when we have no island, and, more specifically, whether it displays a similar acceptability pattern to P-stranding under sluicing compared to P-pied-piping; but also b) how islands and P-stranding interact with each other under sluicing and overtly. With respect to the second point, in particular, it would be interesting to see if overtly island extraction is at floor level, regardless of P-stranding, or if the two have an additive effect on acceptability deterioration; and under sluicing, it would be interesting to see if a) P-stranding out of an island is as acceptable as P-pied-piping without an island, which would be expected if the two are simply islands to be ameliorated across the board[1], or if b) P-stranding behaves differently to regular islands. The results of this experiment showed main effects of both Sluicing and P-stranding as well as an interaction of the two, with sluicing conditions being overall significantly more acceptable than overt conditions; P-stranding being significantly less acceptable than P-pied-piping; and the difference between P-stranding overtly and under sluicing being significantly larger compared to the same difference for P-pied-piping. The important thing to observe here is that the sluicing results of Experiment 2 directly replicated those of Experiment 1, with P-stranding under sluicing still being significantly less acceptable than P-pied-piping, albeit much more acceptable than overtly. A comparative analysis of the two studies' results using Experiment (Experiment 1 vs. Experiment 2) as a between-subjects factor showed that, indeed, there was no significant difference between the two in terms of their sluicing condition results[2]. This comparison represents the first experimental data in the literature to show what island amelioration under sluicing really looks like in terms of acceptability. Based on the results of these two studies we can observe that a) island amelioration under sluicing does not appear to have any effect on sluicing acceptability ratings, with P-pied-piping with or without an island under sluicing obtaining the exact same ratings; and b) this amelioration presents a different pattern to the amelioration of P-stranding. Namely, regular island extraction (or evasion) under sluicing does not result in a deterioration in acceptability ratings compared to when no island is involved, i.e. sluicing completely ameliorates islands; whereas P-stranding under sluicing does appear to result in a significant deterioration in acceptability compared to P-pied-piping, i.e. sluicing does not completely ameliorate P-stranding. Together, the results of Experiment 1 and 2 indicate that PPs do not appear to behave the same way as other, regular islands under sluicing.

To gain more confidence in this result, however, whilst also practically excluding the possibility of a cleft structure being the alternative source for P-stranding under sluicing in Greek - despite the obvious counter-indicative factor of the remnant's case not matching that of a cleft pivot - Experiment 3 examined P-stranding in the context of contrastive sluicing. Specifically, it has been argued several times in the past that contrastive focus blocks the amelioration of regular islands under sluicing (M. Barros et al., 2014; Griffiths & Lipták, 2014; Merchant, 2008). Simultaneously, contrast sluicing or 'else' modification is also disallowed with copulas cross-linguistically. As such, if the acceptability of P-stranding under sluicing in Greek is in any way linked to either an alternative, cleft source, or to its being simply another PF-island for this language, then we would anticipate contrast sluicing to block P-stranding. In other words, we would anticipate no significant difference any more between P-stranding overtly and under sluicing. The acceptability judgements

[1] Although we know from Experiment 1 that this cannot be the case, as we found that P-stranding is *not* as acceptable as P-pied-piping even outside of the context of islands.

[2] Only the case-matching data were used from Experiment 1 for the purposes of this comparison.

results collected for this study were very similar to those of the previous two experiments. Specifically, results showed a significant main effect of Sluicing, with sluicing conditions being more highly rated than overt conditions; a main effect of P-stranding, with P-pied-piping conditions being more highly rated than P-stranding conditions; and a significant interaction between the two factors, with P-stranding being more acceptable under sluicing compared to overtly, with no such difference for P-pied-piping. These findings showed that, despite the restrictive nature of contrastive sluicing, the behaviour of P-stranding overtly and under sluicing was exactly the same as in the previous two studies. Together, the results of these three experiments allow us to gain more confidence in the hypothesis that P-stranding is not a form of PF-island in Greek, nor can it be explained via an alternative, cleft source.

At the end of these three experiments, we are left with an intriguing set of results which no existing theory appears able to explain. On the one hand, movement-based accounts fail to explain how P-stranding can possibly be allowed under sluicing without resorting to either an alternative source or PPs being PF-islands in Greek; whilst non-movement-based accounts, on the other hand, fail to account for the significant difference we found between P-stranding and P-pied-piping under sluicing, if no structure is present at the *e*-site. Simultaneously, however, given the lack of existing controlled datasets on P-stranding in other languages, it is also unclear whether our results are particular to Greek, meaning we should focus our explanation on its unique properties, or whether we might anticipate similar results in *all* overtly non-P-stranding languages once put similarly to the - adequately powered and controlled - test. To discover just how generalisable these results are, therefore, an additional experiment was run, replicating the original Greek acceptability judgement design, this time in German. The factors of P-stranding, Sluicing and remnant Case-matching were, thus, once again crossed. The results of this web-based acceptability study showed results remarkably close to those of the original Greek experiment. Specifically, we found a significant main effect of all three factors, with sluicing conditions being overall more highly rated than their overt counterparts; P-stranding being significantly worse than P-pied-piping; and case-matching being overall better than case-mismatching. Importantly, just as in the original, we also found significant two-way and three-way interactions, with P-stranding being significantly more acceptable under sluicing compared to overtly, although not as acceptable as P-pied-piping. Interestingly, and slightly differently from Greek, case-mismatching was also significantly better under sluicing compared to overtly, with the case-mismatching P-stranding condition being as acceptable as its case-matching counterpart; whilst the overt case-mismatching P-stranding condition was still the least acceptable condition of all.

At the end of all these four experiments there are a number of conclusions that can be quite confidently drawn. First of all, as expected of two overtly non-P-stranding languages, P-stranding overtly in both Greek and German is unacceptable. The same is true for case-mismatching. However, contrary to the PSG, P-stranding in both languages becomes significantly more acceptable under sluicing, although still being significantly less acceptable than P-pied-piping. Furthermore, this is true whether examined in the context of islands or contrast sluicing, at least insofar as Greek is concerned. This last result, together with the different acceptability profile of P-stranding and regular island amelioration under sluicing, indicates that PPs do not appear to simply be another PF-island. The possibility of an alternative, cleft or resumption source is also excluded for both Greek and German. As a whole, this behaviour comes as a surprise to all existing theories of sluicing, particularly when accompanied by the finding that P-stranding with case-mismatching in German sluicing is just as acceptable as P-stranding with case-matching, given all theories require case-matching of remnant and correlate for one reason or another. Furthermore, the inclusion in all four experiments of the factor of *wh*-XP complexity in the experiment design and analyses showed no significant effect of this factor on either P-stranding or case-mismatching acceptability under sluicing. This particular result indicates that a sentence processing approach based on information complexity, following either Ariel (2014) or Hofmeister (2011), such as Nykiel's (2013) account of Polish P-stranding, does not successfully capture these data either. Finally, we can conclude that the amelioration of P-stranding under sluicing is not uniquely tied Greek, but appears to be true in other overtly non-P-stranding languages as well.

At the end of Chapter 2, we are thus left with a baffling mystery as to how to explain these cross-linguistic data. This is particularly true if we take into consideration the only other powerful and controlled existing dataset in a PSG-defiant language, namely the Saudi Arabic (SA) data presented in AlShaalan and Abels (2019), collected following the same processes as in this thesis. In this paper, P-stranding under sluicing in SA was posited to have two sources; one involving resumption; and one presumably involving P-stranding, given resumption is blocked. When the alternative source involving resumption was allowed, P-stranding under sluicing was found to be significantly more acceptable than overtly, and just as acceptable as P-pied-piping. On the other hand, when this alternative source was blocked, P-stranding under sluicing was still significantly more acceptable than overtly, but was now also significantly less acceptable than P-pied-piping. If we are to overcome the limitations of the existing literature and attempt to find a coherent and cohesive explanation of the behaviour of P-stranding under sluicing cross-linguistically, then the data from all three of these languages must be taken into consideration.

This brings us to Chapter 3, where a novel inter-disciplinary approach is proposed to not only explain these data, but also generate testable predictions for all other languages, overtly P-stranding or not. This account combines a theoretical syntax approach, in the form of a PF-Deletion Hypothesis, together with a sentence - and more generally information - processing approach. In accordance with a PF-Deletion Hypothesis, some form of structure is posited at the site of ellipsis which exists in a relationship of mutual semantic entailment with the antecedent. When the remnant behaves as though it is a part of such a structure, then no further processing costs are incurred for the parser to incorporate it, with the transmitter's intended message being clearly conveyed. However, when a remnant is encountered that does not behave as though part of such structure, such as a case-mismatching or P-stranding remnant (in an overtly non-P-stranding language), then the parser engages in a process of probabilistic estimation to gauge how likely it is that the transmitter of the message meant this unacceptable structure as is vs. an alternative, slightly edited but acceptable one. This calculation incurs a series of processing costs which are then translated to differences in acceptability.

More specifically, this approach draws on concepts from the field of information theory (J. Hale, 2001; Levy, 2008a, 2008b; Shannon, 1948) and their existing applications in linguistics (Gibson et al., 2013). At the beginning of Chapter 3, the concept of anticipation or prediction is introduced, with experimental evidence demonstrating that the parser has a tendency to anticipate upcoming linguistic material in advance of bottom-up confirmatory input. This anticipation has been argued to be guided by the frequency of occurence of different structures (J. Hale, 2001), but also context (Shannon, 1948), with this context involving both the sentential context, but also the wider, extra-sentential context of world knowledge (Levy, 2008a). Based on this pre-existing knowledge, as a sentence is unfolding in front of us, as comprehenders it is possible for us to predict what the most likely incoming message is going to be before seeing or hearing it in its entirety. The more likely the message, the more expected, and therefore the easier it will be to parse; the less likely, the less expected or the more surprising, and thereby the harder to parse (Levy, 2008a). One of the reasons the parser may use such prediction techniques is in order to attain the highest efficiency and speed in communication, by constantly anticipating the most likely outcome and planning for it. However, this prediction is also argued to be used to overcome a very important ubiquitous problem in language - and more broadly information - communication, and that is the presence of noise during message transfer (Shannon, 1948). Specifically, although every communicator sets out with the intention to successfully convey a message in the most ideal and efficient way, very frequently this has to be conveyed through a noisy medium. Restricting our attention to linguistic message transfer in particular, this noise can be literal and external to the communicators, such as background music playing in a bar or other people chatting around us; or it can be internal to the speaker or comprehender. For instance, a source of noise which is internal to the speaker may be a speech impediment or a poor grasp of the language as a second language learner, each of which can affect how well-formed and grammatical his or her utterances are. These same sources of metaphorical noise can also be internal to the comprehender, leading him or her to not accurately perceive or dissect the message. Despite the omnipresence of such noise sources, however, thanks to our ability to use pre-existing knowledge to anticipate and parse upcoming material, we are (mostly) able to avoid poor message transfer or total communication breakdown.

An elegant model reframing Shannon's original (1948) communication model and capturing how we may avoid such errors during daily communication is presented by Gibson et al. (2013), repeated below in (5.1). This model forms the basis of the noisy channel approach to sluicing I take here.

$$P(s_1|s_2) \propto P(s_1)(Ps_1 \rightarrow s_2) \tag{5.1}$$

This is a simplified Bayesian probabilistic model where s_1 is the intended sentence or message (for ease, these are considered the same thing here) on behalf of the producer and s_2 is the perceived message on behalf of the comprehender. Specifically, this model states that the posterior probability of the intended message (s_1), given the perceived message (s_2), is proportional to the prior probability of s_1 multiplied by the probability that s_1 could be corrupted into s_2. This probability has been argued to be measured in terms of the number and types of string edits required for the two to coincide (Gibson et al., 2013; Levy, 2008b). In other words, the parser is considered to operate as a Bayesian probabilistic machine, generating hypotheses about incoming messages based on a) the likelihood of s_1 in and of itself given the sentential and extra-sentential context it appears in, as well as b) the likelihood that this s_1 could be altered or corrupted into s_2, again given the context it appears in. Aside from the fact that these two factors make up the posterior probability of s_1, what this equation also tells us is that both of these factors are tied to their surrounding narrower and wider context, along with how much noise this context contains. When no type of noise is present, if a completely acceptable, plausible and likely sentence given the sentential context is perceived, with no string edits required, then the parser will accept this message as is and afford it a high posterior likelihood. If an unacceptable message is perceived, however, the parser has two options: either accept s_1 as is, with no string edits, which has a low prior likelihood; or infer a different, string-edited version of s_1 with a much higher prior likelihood. In the first case, the low prior likelihood will be multiplied by the high likelihood of no string edits. In the second case, the much higher prior likelihood will be multiplied by the likelihood of 1 string edit, in itself slightly lower than the likelihood of no string edits. As a result, the second option will end up with a higher posterior probability, making it a more attractive option for the parser.

However, as Gibson et al. (2013) also showed with a series of experiments, this calculation process is affected by the amount of noise contextually present. This noise can make implausible or unacceptable messages a priori more likely to occur or it can make string edited versions of the intended messages more likely to occur. In a situation where many implausible messages occur, this will make the parser more likely to accept such an implausible message. On the one hand, therefore, it is possible to raise the prior likelihood of an unacceptable message by manipulating the number of such unacceptable or implausible messages contextually. I will term this type of noise ‘implausible message noise’. On the other hand, it is also possible to make the parser more likely to assume that the s_2 perceived is not in fact the intended s_1, but rather that a string-edited version of s_2 should be inferred instead. In other words, it is also possible to raise the probability of corruption of the message, by increasing the amount of syntactic errors contextually. In a situation where many such errors occur, such as when speaking to a second-language speaker, the parser will be more likely to assume a string-edited version of s_2 was intended. I will term this type of noise ‘string-edit noise’. The likelihood of corruption, i.e. of string edits, is further affected by the number and type of string edits. The larger the number of string edits required to reach an acceptable version of s_1, the smaller the likelihood of corruption. With respect to the type of edits, some argue that a deletion (i.e. that something has been lost during message transmission) is more likely than an insertion (i.e. that something has been erroneously inserted during transmission) (Gibson et al., 2013), whereas others consider the two to be equally likely (Levy, 2008b).

Based on this idea of probabilistic determination through a language channel that is more or less noisy, depending on context, I argue that the acceptability judgements found in all the experiments presented in Chapter 2 are off-line reflections of the processing costs associated with the parser undergoing this process of inferring the intended message with the highest posterior probability. Sentences with sluices contain a substantial amount of missing material which must be inferred through imperfect memory mechanisms in conjunction with our prior expectations. This leads me to argue that sluicing represents a naturally noisier channel of communication compared to full overtly expressed structures, where no such inference is necessary. As such, the parser is more likely to accept that a message perceived via sluicing might require string edits in order to reach an acceptable intended message, compared to a full overt message. Consequently, when the parser encounters a sluice remnant missing a P in an overtly non-P-stranding language, the already evident partial deletion of the signal together with these less-than-perfect memory mechanisms may lead the parser to converge on an intended representation that includes a pied-piped P which has simply been deleted. The cost associated with making this 1 string edit to infer the acceptable intended message is reflected in a lower acceptability rating associated with the ‘P-stranded’ remnant compared to a remnant with an overtly expressed pied-piped P. However, given the noisy channel of sluicing, this string edit is considered more likely than the omission of a P would be in an equivalent overt structure, where noise is comparatively reduced thanks to memory mechanisms not being required to access missing material. Given these instances of sluicing have typically been considered instances of covert P-stranding, in my experiments I do not compare P-stranding under sluicing to overt P-omission, but rather to overt P-stranding, which is unacceptable in these overtly non-P-stranding languages. As such, a significant difference in acceptability is observed between a) P-stranding overtly; b) P-'stranding' under sluicing; c) P-pied-piping overtly and under sluicing.

This is what is assumed to happen if a P-stranding remnant is observed in an overtly non-P-stranding language which has no alternative source available for such a remnant. If, however, we are in a situation where an alternative interpretation of the perceived message is available, i.e. an interpretation where it is no longer unacceptable, the model's predictions are slightly different. If, for instance, a copular or resumptive source can be inferred instead of a P-stranding one, then the parser need not go through the process of calculating and inferring string edited versions of this message. Instead, it can simply accept the message version which is acceptable with no string edits required. Given no such edits are necessary, such 'P-stranded' sluices should also not require any processing costs to parse, with their acceptability rating simply being tied to that of this alternative acceptable intended message. However, when such an alternative is somehow blocked, then the parser must go through the same process as though this message does not exist. In other words, it must infer a string-edited version of the perceived message, leading to a lower acceptability rating. This leads to the pattern observed, for instance, in SA.

Based on the above logic, a straightforwardly testable set of cross-linguistic predictions are made, repeated below in Table 5.1, for how the parser should behave when encountering a remnant without a P.

Table 5.1: Cross-linguistic P-Stranding Predictions for Non-P-Stranding Languages

'The Maria danced with someone, but not know ...

			Acceptability*		
	Overt Structure	*e*-site	L1	L2	L3
Overt	a) ... *with who she danced.'*	N/A	6/7	6/7	6/7
	b) ... *who she danced with.'*	N/A	2/7	2/7	6/7
	c) ... *who it was with which she danced.'*	N/A	6/7	2/7	6/7
Sluicing	d) ... *with who*	[she danced].'	6/7	6/7	6/7
	e) ... *who*	[she danced with].'	4/7	4/7	6/7
	f) ... *who*	[it was with which she danced].'	6/7	N/A	6/7

* *Using a scale of 1 - 7, with 1 being completely unacceptable and 7 being completely acceptable.*

Specifically, for overtly P-stranding languages (e.g. L3), such as English, the parser should be able accept the remnant as is. That is to say, with no other processing costs being taken into consideration here, the posterior likelihood of P-stranding should simply be tied to the prior likelihood of P-stranding overtly in embedded clauses, with no string edits being necessary. Following the concept that similar processing costs lead to similar acceptability ratings, we would thus expect P-stranding overtly and under sluicing to receive similar acceptability ratings in such a language.

When the language in question is overtly non-P-stranding, then the behaviour of the parser rests on the availability or not of an alternative acceptable structure to infer at the e-site. When such an alternative is available with no further string edits being necessary, such as a copular structure in Brazilian Portuguese, then what appears as P-stranding under sluicing should simply be interpreted as an instance of this alternative structure (e.g. L1). Following the same logic as above, the acceptability of this apparent P-stranding sluice should be equal or similar to the acceptability of this alternative structure overtly[3].

However, when such an alternative structure does not exist or is somehow blocked (e.g. L2), then the parser must calculate the s_1 with the highest posterior likelihood, based on the prior likelihoods of various s_1's along with the number of string edits required to reach them from the perceived 'P-stranded' s_2. Theoretically, the closest s_1 to the perceived s_2 in this case should always be an otherwise identical message which simply involves P-pied-piping, with the string edit necessary for s_1 and s_2 to coincide being simply one deletion. Furthermore, assuming that a) deletions are likely string edits in a noisy channel and b) sluicing is a naturally noisy channel, the likelihood of corruption will be calculated as relatively high, compared to e.g. the same corruption occurring in a full overt clause. As such, multiplying the high prior likelihood of an s_1 containing P-pied-piping by the high likelihood of corruption in this context, the parser arrives at a high posterior likelihood for P-pied-piping being the intended s_1 based on the perceived s_2. The cost of inferring this string-edited s_1 however, will be larger than when no string edits are required. As such, this P-less remnant will be more costly to parse and therefore less acceptable compared to such a remnant in a) an overtly P-stranding language and b) a situation where an alternative structure is available. It should also be more costly compared to a P-pied-piping remnant, i.e. a remnant which requires no string edits.

Based on this logic, Greek and German as overtly non-P-stranding languages should naturally fall into category L2, whereas SA should behave as an L1 language when an alternative structure is available and as an L2 one when such a structure is blocked. This is exactly the pattern of ratings we found in Chapter 2 and AlShaalan and Abels (2019) found in their experiments.

In keeping with this noisy channel account of sluicing, one would anticipate that the larger cost of parsing P-stranding remnants compared to P-pied-piping ones in languages such as Greek and German should also manifest itself behaviourally in a difference be-

[3]Given we do not have any acceptability ratings from such languages other than in Saudi Arabic, to keep things simpler for the reader, I have simply assumed for a language of type L1 a given rating associated with the overt alternative structure in (c) (e.g. 6/7) and consequently the sluice supposedly containing this structure as an alternative sluice source (f). This rating is simply for illustration purposes, however. There could theoretically be other languages where this overt alternative is rated as more or less acceptable (e.g. 5/7 or 6.5/7) compared to embedded *wh*-questions with P-pied-piping (a). In each of those situations I assume that the sluice rating in (f) will be very similar to this overt rating in (c).

tween the two in online measures. Moving forward, therefore, this was the next logical step in collecting evidence for this account, i.e. to gather online evidence suggesting that P-stranding is indeed costlier to process under sluicing compared to P-pied-piping in overtly non-P-stranding languages. This brings us to Chapter 4. With this goal in mind, a review of the existing literature identified a specific study targeting sluicing predictions ahead of bottom-up confirmatory input, using the method of SPR (Yoshida et al., 2013), which was considered an ideal candidate to measure these processing differences for our intents and purposes. Specifically, if we do indeed predict sluicing in advance of bottom-up input in the environments this study claimed we do, then when such a sluice projection is generated it should be possible to measure the difference in processing times when the parser encountered a P-less sluice remnant vs. a P-pied-piping one in Greek and German.

The original study was run in English, with the claim of a default sluice prediction resting on our reaction to a reflexive pronoun contained within this supposed sluice projection and constrained therein by the rules of binding. More specifically, when a sluice was supposedly being projected, the study showed that we react when the gender of this reflexive pronoun does not fit this sluice projection, with participants exhibiting a GMME at the reflexive, but with no such reaction when the reflexive's gender fit the projection. On the other hand, when a sluice projection was blocked, participants no longer reacted to the reflexive pronoun's gender.

Before moving forward in manipulating the design to measure the cost of P-stranding in Greek, given the original study was conducted in another language, it was first necessary to establish that its design and results could indeed be replicated in Greek. A replication study of equal power to the original was thus run with Greek native speakers. However, the original findings were not replicated, with this study instead showing an intriguing difference in results. Specifically, the Greek results showed a GMME at the reflexive regardless of sluice predictability. This lead us to question whether the original study results were accurately interpreted or whether they may have been (partially) driven by other factors. Specifically, the two problems identified were, firstly, that the reflexive used in both studies was contained within a representational NP; however, anaphors contained within such NPs have been argued to be exempt from the rules of binding; as such, the GMME observed at the reflexive in these studies may not have been due to a sluice being projected, but rather due to the reflexive itself generating a search for its antecedent and failing to find a suitable gender-matching candidate. The reason this GMME was not present in the original study results when a sluice was not predicted is tied to the second study issue identified: a local attachment ambiguity very near the reflexive in the English study. Specifically, this local PP attachment ambiguity, which was present only in the sluice non-compatible conditions, may have derailed the parser, making the reflexive gender mismatch go unnoticed. Importantly, this ambiguity was not present in the Greek replication, hence the GMME always being noticed in the Greek study. Together, these two issues meant that the original study findings did not necessarily indicate a sluice being projected.

That having been said, this did not mean that the original claim of sluicing being projected ahead of time in the correct environment was actually incorrect. The possibility still existed that we may have a tendency to predict sluices ahead of time if the context allowed them, we simply had to find a different way to measure this tendency, whilst removing potential confounding factors. If we can show this to be true, then we could use this modified study version to measure the effects of P-stranding as originally planned. With this goal in mind, as well as clarifying the original study results, a modified version of the study, excluding both reflexive and local attachment ambiguity, was run in English to gather evidence for or against such a sluice tendency. The concept behind this study design was that if we have an innate tendency to project sluicing wherever possible, then when such a sluice is shown to not be the correct continuation we should find a reanalysis slow-down in RTs. On the other hand, when a sluice continuation is blocked, then no such slow-down should be present. Unfortunately, however, this new study also found a null result, providing no evidence for an innate preference for sluicing.

Given all three of these SPR studies were based on the presence or absence of a pied-piped P to either block or allow a sluice prediction, respectively, we also took the opportunity in this last study to investigate how easily participants incorporate pied-piped Ps, as it has been claimed that P-pied-piping in English embedded clauses may in fact be very hard to parse compared to P-stranding, if not being entirely unacceptable. We, therefore, checked how easy it was to integrate a pied-piped PP when it could be expected (i.e. when the supposed sluice antecedent contained the same PP), vs. when it could not (i.e. when the supposed sluice antecedent did not contain this PP). We found that it was easier to integrate such a PP when it was also contained in the antecedent. Furthermore, down-stream evidence from the sentence point where a sluice prediction would be disambiguated showed no difference between these two conditions, indicating that it was not the case that an integratable P encouraged a sluice prediction whereas a non-integratable one did not. We took this finding as evidence that the difference in parsing the PPs between the two conditions was not due to the P being more easily integrated into some projection, but rather that in the condition where the antecedent contained the same PP, the P was simply lexically primed.

One last piece of evidence for or against a sluice prediction was gathered through the use of a SC task, as was also used to supplement the Yoshida et al. (2013) study. Specifically, the same stimuli as in the last, simplified Enligsh SPR task were used, truncated before a *wh*-XP was encountered, i.e. before a sluice was specifically cued. The idea here was to see what sort of continuation participants preferred, with continuations being classified as sluices, parallel structures, clefts or simply ‘other’. If we have a tendency to project sluicing (or clefts or parallel structures, for that matter), then we should find a clear preference for this sort of continuation over the others. Instead, what we found was that 'other' types of continuation (with a Mean of 75.06% across all conditions) were significantly more frequent than sluices (Mean: 12.27%), parallel structures (Mean: 5.137%) or clefts (Mean: 6.135%), regardless of condition. This result once again provided further evidence against sluices, or ellipsis more generally, being innately preferred wherever possible compared to other structures. This SC study also allowed us to see whether a pied-piped P can serve to block a sluice continuation, as was the key assumption in the three SPR studies in order to theoretically be able to distinguish between sluice-compatible vs. sluice non-compatible structures. Instead, however, the SC study actually showed that these pied-piped Ps did not serve to block sluice projections, instead seeming to affect only parallel or cleft structures as potential continuations. As such, despite the theoretical soundness of the SPR design manipulations, it appears as though they did not necessarily create the desired comparison between different conditions. As a result, at the end of all these experiments, we were forced to conclude that a) the original SPR study results were likely driven by spurious factors and b) we do not have an innate preference for sluice projections. This also foiled our intent to use this design in future studies to measure the processing cost of P-pied-piped vs. P-stranded sluice remnants in Greek and other such languages. Due to the temporal constraints of this PhD, at the end of these studies we were, unfortunately, left without enough time to design and implement a suitable alternative paradigm to target potential processing costs associated with P-stranding vs. P-pied-piping in overtly non-P-stranding languages. However, we did conclusively show that relying on the presence or not of a P to make the distinction between sluice compatibility and sluice non-compatibility of stimuli in a research design is not a reliable mechanism. When we discuss future directions for this research in the next section, other potential designs which could be used to provide evidence for the noisy channel hypothesis argued for here will be proposed and examined in more detail.

To briefly recapitulate, over the course of these four chapters, we have introduced the concept of sluicing, a form of ellipsis, as well as the intriguing phenomenon of P-stranding under sluicing in certain languages. Specifically, we identified a large gap in the existing literature with respect to a) the existence of adequately powered and controlled experimental studies to support P-stranding being allowed or not under sluicing in overtly non-P-stranding languages, as well as b) a cohesive cross-linguistic explanation for this behaviour. As a result, we provided the first experimental data of its kind showing that P-stranding is indeed allowed under sluicing in two overtly non-P-stranding languages, namely Greek and German. What is more, this P-stranding is significantly better under sluicing compared to overtly, however is still significantly less acceptable than P-pied-piping. This intriguing result was explained through a novel inter-disciplinary account of sluicing, with sluicing theoretically being explained via a PF-Deletion Hypothesis, with silent structure existing in the *e*-site in a relationship of mutual semantic entailment with the antecedent, but with sluicing being considered a naturally noisy channel for the parser. The main idea behind this account is that the parser behaves as a Bayesian estimation calculator, weighing up the prior likelihood of an intended message together with the likelihood that this intended message may have been corrupted into the perceived one, whilst taking into consideration the surrounding sentential and extra-sentential context, in order to reach the posterior probability of this intended message. Based on this concept, sluicing, due to its lack of overtly expressed material and thereby reliance on imperfect memory mechanisms and contextual knowledge to fill in this missing material, is considered an inherently noisier context compared to an equivalent non-sluicing, overtly expressed utterance. This noisier context is argued to make the parser more likely to accommodate a less-than-perfect perceived message and to infer, instead, an alternative, slightly edited but acceptable intended message. Perceived P-stranding under sluicing in overtly non-P-stranding languages, can thus be inferred to be an instance of P-pied-piping where the P has been deleted. The process of having to infer this deleted P carries with it a processing cost, translating to P-stranding under sluicing being less acceptable in these languages compared to P-pied-piping. Importantly, based on this analysis, a series of cross-linguistic predictions are further made, predicting that *all* overtly non-P-stranding languages should behave similarly to Greek and German under sluicing. The exact acceptability ratings should be modulated by the existence or not of an alternative inferable structure in what appear to be P-stranding instances under sluicing. Based on these off-line data, further online predictions are made with respect to the cost of P-stranding vs. P-pied-piping in these languages, however the process chosen to target these costs in the end proved to be unsuitable. As such these on-line costs were not able to be demonstrated during the short time-course of this PhD.

5.2 Limitations and Future Directions

Given the time constraints inherent in conducting a PhD, as well as experimental setbacks that frequently occur, naturally there are always limitations to doctoral theses, with this one being no exception.

One of the first limitations here is that it was not possible to run a study where we could measure the online processing costs potentially associated with P-stranding and case-mismatching compared to P-pied-piping and case-matching in overtly case-marking non-P-stranding languages. This was due to the original method selected to target these processes, namely by manipulating the form of the remnant in supposed sluice projections, proving to be inadequate. Specifically, when we embarked upon a process of replicating a SPR study by Yoshida et al. (2013) which claimed that we inherently project sluicing when presented with the opportunity, it was found that this was not the case and that the original study results were likely driven by separate, confounding factors. As such, the time spent on this process, whilst very enlightening with respect to the question of whether we have an innate tendency to project sluicing, did not allow for alternatively designed attempts to target these processing costs. Given our explanation of the differences found in the acceptability results of Chapter 2 are based on such, presumed, processing costs, practically showing these costs via properly conducted experiments is a key next step. As such, identifying a way to measure these costs online should be one of the main aims of this research moving forwards. One possible way of doing this is examined a little later on.

A second limitation, related to this first one, is that, once again, due to time limitations, we were unable to test this noisy channel theory in more depth. That is to say, taking into consideration the idea behind the Bayesian equation in (5.1) and Gibson et al.'s further experimental manipulations, the posterior probability of any intended message, or in other words the likelihood that the parser will accept this intended message, is assumed to be the product of two variables, the prior likelihood of the intended message, and the likelihood of this message being corrupted into the perceived one. These likelihoods, however, can both be influenced by the amount of contextual noise present. Specifically, it is possible to increase the environmental noise present in order to increase the likelihood of string edits having occurred to the intended message; this is the type of noise that we previously termed string-edit noise. This string-edit noise is also the type of noise I argue is inherently present in the context of sluicing to a larger extent than in the context of fully overtly expressed structure, thanks to part of the message already being missing overtly and requiring 'filling in'. If this is true, then moving forward there are several further manipulations which can be made to provide further evidence for and clarify this behaviour.

The first two manipulations relate to whether sluicing indeed presents a context with a larger amount of this string-edit noise compared to a context containing equivalent overtly expressed material, as is the basic idea behind the approach advocated for here. To examine this, the first experimental manipulation would be to compare the acceptability of P-stranding under sluicing to that of P-deletion overtly, as the experiments presented here only compare the former to the acceptability of P-stranding overtly. If the parser is more amenable to accepting deletions under sluicing, then we should find a clear difference in acceptability between these conditions, with P-deletion overtly being significantly less acceptable than P-stranding under sluicing. The second experimental step in this direction would be to focus on P-deletion overtly and manipulate the amount of string-edit noise present in the fillers, by inserting syntactic errors, to try to push the acceptability of P-deletion overtly more towards the acceptability of P-stranding under sluicing. If sluicing is simply a noisier context, then following Gibson et al. (2013), it should be possible to make P-deletion overtly as acceptable as under sluicing simply by creating a noisier environment in terms of this type of noise. Both of these manipulations together would provide further evidence for this account of sluicing being on the correct path.

The next manipulations to be made should target the type of edits required to infer an acceptable intended message both under sluicing and overtly. Specifically, as mentioned previously, Gibson et al. (2013) found evidence that string edits in the form of insertions were less acceptable compared to those in the form of deletions. To provide evidence for a) sluicing being a noisy channel and b) P-stranding being a form of deletion, we should, therefore, look at how inserting spurious Ps compares to deleting these Ps in these two contexts. To this end, we could cross the factors of Context (Sluicing vs. Overt Continuation) and Edit Type (No edit vs. Deletion vs. Insertion) to create the table of conditions in 5.2. In this way, stimuli containing additional Ps and stimuli containing no Ps would be compared to each other and to regular P-pied-piping, both under sluicing and overtly. If our assumptions are correct, we would thus predict a significant main effect of Edit Type on acceptability ratings, with no edits being significantly better than deletions, which in turn should be significantly better than insertions. We would also predict a significant interaction of the two factors, with no edits being equally acceptable under sluicing and overtly, but with both deletions and insertions being more acceptable under sluicing than overtly, with deletions, further, being better than insertions. As in the previous set of manipulations described above, these string edits could also be examined overtly in the presence of more vs. less filler string edit noise.

Another interesting question to address would be whether this acceptable deletion under sluicing extends to other linguistic features and/or elements, as a general noisy channel approach would predict, or whether there is something particular to Ps which make them easier to infer when missing. To examine this, one could target a different element for deletion under sluicing, such as, for instance, a missing Determiner, or to target a linguistic feature, perhaps a missing number marking. If there is something inherently special to Ps, then we should find a different set of results to the above manipulations when targeting each of these other elements.

Furthermore, moving on from P-stranding to the interesting data we collected on case-mismatching, it would also be a good idea to observe how the acceptability of case-mismatching overtly and under sluicing compares to the results in Table 5.2. Specifically, in order to keep our interpretations inline with existing noisy channel literature, in this thesis string edits were considered to only be made up of deletions and insertions. As such, case-mismatching was considered a form of deletion followed by insertion, or vice versa. If this is indeed true, then according to the Bayesian size principle, case-mismatching overtly and under sluicing should be considerably less acceptable than P-deletion, P-insertion and no edits in both of these contexts, given it represents a larger number string edits required to reach the intended message. If, however, substitution is simply a form of insertion or represents its own form of single string edit, then we should find that it behaves similarly to these single string

Table 5.2: Follow-up Noisy Channel Study Conditions

(152) Chthes to vradi i Maria chorepse me
Yesterday Det.N.ACC *night*.N.ACC *Det*.F.NOM *Maria*.F.NOM *danced*.3SG *with*
kapjon, alla de thimame…
someone.M.ACC *but* NEG *remember*.1SG …
'Last night, Maria danced with someone, but I don't remember…'

	Condition	Example
1	Overt Continuation No Edit	… me pjon chorepse. … *with who*.M.ACC *danced*.3SG '… with whom she danced.'
2	Overt Continuation, Insertion	… me se pjon chorepse. … *with to who*.M.ACC *danced*.3SG '*… with to whom she danced.'
3	Overt Continuation, Deletion	… pjon chorepse. … *who*.M.ACC *danced*.3SG '… whom she danced.'
4	Sluicing No Edit	… me pjon. … *with who*.M.ACC '… with whom.'
5	Sluicing Insertion	… me se pjon. … *with to who*.M.ACC '*… with to whom.'
6	Sluicing Deletion	… pjon. … *who*.M.ACC '… whom.'

edits (even though we have already observed in our experiments that case-mismatching does not appear to behave exactly like P-stranding). As a side-note here on the subject of case-mismatching, I would also like to remind the reader of the interesting difference we found between Greek and German with respect to case-mismatching under sluicing. Specifically, we found that German speakers were more amenable to case-mismatching under sluicing when coupled with P-stranding, as much so in fact as case-matching, whereas Greek speakers did not present the same pattern. The explanation proposed then was that cleft continuations in the context of a PP antecedent may be more common in German compared to Greek, thereby encouraging a cleft alternative structure to be a close inferable intended message for German speakers, but not Greek ones. To delve deeper into the cause of this difference, therefore, a follow-up question which would be very interesting to conduct would be to examine how German participants react to case mismatching with a non-cleft case, e.g. when instead of Dative, the remnant appears in Accusative or Genitive. If the difference we found between the languages for case-mismatching is due to cleft or copular structures being more readily available as alternative structures at the *e*-site for German but not as much for Greek, then when we block this copular alternative, we would anticipate German participants to rate case-mismatching with Accusative or Genitive as much worse than Nominative. If, on the other hand, they are willing to accept case mismatching even with these non-cleft cases, then that would indicate that the parser is indeed weighing up probabilities and accommodating for the incorrect case in the same way that it does for P-stranding. A similar cleft-blocking situation could be achieved in German by utilising the context of contrast sluicing, similarly to our Greek Experiment 3, given contrastive focus is considered by default to disallow copular continuations.

Finally, with respect to the behaviour of sluicing as a noisy channel and bearing in mind, once again, the Bayesian equation in (5.1), aside from the above manipulations which all target the amount of string-edit noise present and thereby the possibility of corruption of the intended message into the perceived one, it is also possible to increase the environmental noise in a different way. Specifically, it is possible to increase the noise in order to target and raise the prior likelihood of an implausible or unacceptable message being the intended one, without further string edits being necessary. This is the type of noise I previously termed message plausibility noise. If sluicing is indeed a noisy channel and P-stranding is indeed a form of string edit, then we should be able to manipulate the message plausibility noise present to make P-stranding become as acceptable as P-pied-piping under sluicing and overtly. The way that Gibson et al. (2013) manipulated this factor was to increase the number of implausible or unacceptable filler items presented to participants, thereby making them more likely to accept a perceived unacceptable stimulus as the intended one. As such we could possibly run a similar study to the one detailed above in Table 5.2, only this time increasing the amount of implausible message noise in the fillers and measuring how this affects acceptability ratings for each condition.

Overall, all the above proposed manipulations have targeted the off-line effects of perceived message corruption and intended message inference, depending on the amount and type of noise present. However, in order to target the potential on-line effects of these behaviours, i.e. measure the presumed processing costs associated with P-deletion and/or case-mismatching, as was our original intention in Chapter 4, it may be possible to simply present participants with the same stimuli in an eye-tracking or self-paced reading environment and observe their behaviour when they encounter the 'corrupted' element. For instance, when in Greek the potential sluice antecedent contains a PP-contained indefinite, how do participants react in real time when they encounter a bare *wh*-phrase down-stream compared to a PP-contained *wh*-phrase? If this unexpectedly bare *wh*-phrase is harder to parse, as predicted, it may result in significantly more regressions in an eye-tracking paradigm compared to the PP-contained one; in self-paced reading, on the other hand, the former may result in significantly slower reading times compared to the latter. These measurements of the on-line costs associated with P-deletion could then be compared to the on-line costs associated with a) the same deletion in the presence of more or less noise; b) other forms of feature deletion, such as deletion of number-marking; and c) other types of string edits, such as insertion (e.g. adding unexpected Ps or other features) and/or substitution (e.g. P substitution or case- or other feature-mismatching).

These are simply some of the future directions which can be taken to further clarify both the validity of the noisy channel approach advocated for here, but also the behaviour of the parser with respect to P-deletion more generally. In order to gain more confidence towards any cohesive cross-linguistic account of P-stranding and/or case-mismatching under sluicing, however, it is also of imperative importance to gather further, similar datasets from as many more languages as possible. This is something we intend to do in future follow-up studies, targeting Italian, Spanish, Brazilian Portuguese, English and Czech, to name but a few. Together, all these data will help elucidate the true nature of this phenomenon and answer a decades-old question with respect to how the theoretical and processing literature should explain P-stranding under sluicing moving forward.

5.3 Concluding Contributions

As a concluding remark, regardless of any other contributions this thesis has offered the literature with respect to the particular phenomenon of sluicing, and P-stranding under sluicing more specifically, I believe there are certain additional key things that it makes clearly evident. The first of these is how important large and well-controlled experimental datasets are for the advancement of both theoretical and applied linguistics concepts. Large acceptability judgement studies should be considered the norm to support any sort of cross-linguistic generalisation, as these bring to light subtle differences and gradations in acceptability for what previously have been considered broadly 'acceptable' or 'unacceptable' examples by theoretical linguists and which need to be further accounted for, showing that different phenomena may be a lot more nuanced than originally thought. If explanations of cross-linguistic behaviours are the key aim of theoretical linguistics, then surely the more acceptability data gathered, the better equipped linguists will be to generate such explanations as well as to pre-emptively anticipate and account for behavioural counter-examples which have historically been the bane of any syntactic or semantic theory. From this respect, despite the effort and time associated with gathering and analysing such datasets, in the long run, doing so will actually save time and effort. Furthermore, the design of these experiments should take into consideration existing explanations of this behaviour in order to gather more evidence for or against such explanations, such as was done here, for instance, to investigate the claim that P-stranding under sluicing may simply be a grammatical illusion. Altogether, well-designed and large-enough datasets should prevent regurgitation of the same or slightly altered example stimuli with different acceptability judgements from paper to paper, as seems to currently be the case for a lot of linguistic phenomena.

The second contribution, I believe, is in highlighting the importance of taking into consideration as much data as possible to generate a cross-linguistically applicable explanation for any linguistic phenomenon. Once again, narrower explanations, which are not based on such cross-linguistic datasets, face the much more likely prospect that data from another language will quickly appear to overturn this theory, as we have seen happen repeatedly in the literature when examining just the phenomenon of P-stranding under sluicing. Although it is of course impossible to prevent this sort of thing from happening, and indeed the whole point of scientific research is to gather more data in order to disprove hypotheses, by gathering larger datasets from more languages to begin with, we can at least try and limit this phenomenon from having to happen as often as it currently appears to.

Based on all of the above, I would like to conclude that regardless of whether the explanation provided here for P-stranding under sluicing proves to be correct in the long run or not, I believe it is still of paramount importance to consider the data collected here from both Greek and German, together with the existing dataset on Saudi Arabic and any other one collected in the future from other languages, in order to fully explain this phenomenon without limitations. The same is true for any other linguistic phenomenon one decides to examine in more depth. It is only when we take all such powerful, controlled and detailed cross-linguistic datasets into consideration together moving forward that we will be able to make true progress in the explanation of any phenomenon.

Appendices

Appendix A

Appendix: Chapter 1

A.1 Nykiel (2013) Polish Experiment Design

The first experiment in Nykiel (2013) crossed the factors of P-stranding (P-stranding vs. P-pied-piping) and stimulus type (Overt direct *wh*-question vs. sluicing), presenting participants with four conditions. The first problem regarding the methodology of this study is that the stimuli were presented to participants in the form of a questionnaire. By presenting the stimuli all at once in this way, the researcher allows the participant to compare and contrast them, going backwards and forwards and altering their responses as they read more stimuli. To this end, the types of fillers themselves could also have played a significant role in how participants judged the experimental stimuli, particularly with respect to the less clearly defined conditions in terms of acceptability, i.e. P-stranding/omission; that is to say, more acceptable fillers could have easily pushed the acceptability of P-stranding down and less acceptable ones could have pushed it up. Given there is no mention of the types of fillers used or their grammaticality distribution, this is something we cannot know for sure. Thirdly, not including comprehension questions in the study begs the question of whether participants were indeed reading the stimuli properly and paying adequate attention to consider their data representative.

With respect to the stimuli themselves, it is not stated clearly whether any of the verbs used in the antecedents could potentially take a direct P-less object in this same case or not, regardless of meaning. For instance, if the antecedent verb can pattern both with a PP (e.g. P + NP.ACC) and a bare case-marked remnant (e.g. NP.ACC), with or without the exact same meaning, then an apparently P-stranded remnant could be considered the direct bare object of the antecedent verb, even with a slightly altered meaning. In this case, this would evidently make the 'stranded' remnant more acceptable, since it would not be considered a clear case of P-omission or stranding. The silence of sluicing makes it possible to posit a silent continuation with possibly a slightly altered meaning for the remnant compared to the antecedent correlate. This is, after all, the reasoning for alternative sources in sluicing. The authors would presumably argue that this is being controlled by comparing sluices with apparent P-stranding to overt *wh*-questions with P-stranding, i.e. if the stranded remnant under sluicing could be interpreted as the direct object of the verb, then this would presumably be possible overtly as well. The answer to this, however, is that with overt direct *wh*-questions there is the issue of the overtly expressed stranded P, which does not allow such an alternative interpretation. Based on this distinction, but also the fact that we are dealing with an embedded question for sluicing conditions vs a direct question for overt conditions, the comparison of these two cannot be considered minimal, nor the acceptability of overt P-stranding clearly indicative.

Finally, the types of statistical analyses used for each experiment and the way they are reported is not entirely clear. It is mentioned, for instance, on p. 81 and 83, that 'a mixed-effects regression model' was fitted to the data. Given the DV was a seven-point Likert scale rating, this is evidently not a logistic regression model, but rather a linear mixed effects one. If this is so, then, for instance, in Experiment 2, it is not made clear which factors exactly are being crossed and how they are being reported. A close examination of the example conditions in Nykiel (2013, (28), p.83) and the accompanying results figure (their Fig. 2) indicates that the three factors of *Wh*-Type (Which-NP vs. Wh-XP), Construction type (Overt Cleft vs. Sluice) and P-pied-piping (P-pied-piping vs. P-stranding) were crossed to yield 8 conditions. The reporting of main effects however is confused with interactions and vice versa, with multiple t-values reported with unclear origin (e.g. 'a main effect of Construction type such that clefts with *which*-NP phrases were rated worse than either type of sluicing (*which*-NP sluice: $t = 7.199, p < .00001$; *wh* sluice: $t = 5.098, p < .00001$), and unreliably different than clefts with bare *wh*-phrases ($t = .08$, $p = .94$)[emphasis in original]' a.o.). Additional pairwise t-tests are also run to investigate targeted comparisons, however given each fixed effect factor included in the LME model has only two levels, it is unclear why these additional comparisons are necessary, since significant differences between any pair of interest would appear as a significant main effect and/or a two-way or three-way interaction in the LME output. These t-tests, furthermore, report multiple t-values and degrees of freedom with unclear origin and it is not stated whether they were corrected in any way for multiple comparisons, in itself an issue as this would evidently inflate the likelihood of false positives being reported. These are some examples of how the reporting of results in the paper makes these experimental findings less clear than they could be.

Appendix B

Appendix: Chapter 2

B.1 Experiment 1: P-Stranding under Regular Sluicing (Greek)

B.1.1 Items

B.1.1.1 Experimental Items

To save space only the first item will be presented in all 8 conditions, with all subsequent items being presented only in their Case-matching, Non-Sluicing, P-pied-piping version.

1. a. **Condition 1: Case-Matching, Non-Sluicing, P-pied-piping**
   Λόγω του φορτωμένου ωραρίου της η θεία αναγκάστηκε να παραιτηθεί από κάποιον σύλλογο, αλλά δε μας πληροφόρησε από ποιον σύλλογο αναγκάστηκε να παραιτηθεί.

   b. **Condition 2: Case-Matching, Non-Sluicing, P-stranding**
   Λόγω του φορτωμένου ωραρίου της η θεία αναγκάστηκε να παραιτηθεί από κάποιον σύλλογο, αλλά δε μας πληροφόρησε ποιον σύλλογο αναγκάστηκε να παραιτηθεί από.

   c. **Condition 3: Case-Matching, Sluicing, P-pied-piping**
   Λόγω του φορτωμένου ωραρίου της η θεία αναγκάστηκε να παραιτηθεί από κάποιον σύλλογο, αλλά δε μας πληροφόρησε από ποιον σύλλογο.

   d. **Condition 4: Case-Matching, Sluicing, P-stranding**
   Λόγω του φορτωμένου ωραρίου της η θεία αναγκάστηκε να παραιτηθεί από κάποιον σύλλογο, αλλά δε μας πληροφόρησε ποιον σύλλογο.

   e. **Condition 5: Case-Mismatching, Non-Sluicing, P-pied-piping**
   Λόγω του φορτωμένου ωραρίου της η θεία αναγκάστηκε να παραιτηθεί από κάποιον σύλλογο, αλλά δε μας πληροφόρησε από ποιος σύλλογος αναγκάστηκε να παραιτηθεί.

f. **Condition 6: Case-Mismatching, Non-Sluicing, P-stranding**
Λόγω του φορτωμένου ωραρίου της η θεία αναγκάστηκε να παραιτηθεί από κάποιον σύλλογο, αλλά δε μας πληροφόρησε ποιος σύλλογος αναγκάστηκε να παραιτηθεί από.

g. **Condition 7: Case-Mismatching, Sluicing, P-pied-piping**
Λόγω του φορτωμένου ωραρίου της η θεία αναγκάστηκε να παραιτηθεί από κάποιον σύλλογο, αλλά δε μας πληροφόρησε από ποιος σύλλογος.

h. **Condition 8: Case-Mismatching, Sluicing, P-stranding**
Λόγω του φορτωμένου ωραρίου της η θεία αναγκάστηκε να παραιτηθεί από κάποιον σύλλογο, αλλά δε μας πληροφόρησε ποιος σύλλογος.

2. Εχθές στην πινακοθήκη η μητέρα μου ενθουσιάστηκε με κάποιους διάσημους πίνακες ζωγραφικής, αλλά δεν είδα προσωπικά με ποιους πίνακες ενθουσιάστηκε.

3. Στο προαύλιο του σχολείου η νεαρή μαθήτρια κρυβόταν από κάποιους, αλλά κανείς δεν κατάλαβε από ποιους κρυβόταν.

4. Όλο το πρωί η πωλήτρια επέμενε θυμωμένα στη βοηθό της να μιλήσουν για κάποιον και παραξενευόμουν για ποιον επέμενε να μιλήσουν.

5. Το νέο ντοκιμαντέρ έλεγε πως η επιβίωση των πολικών αρκούδων εξαρτιόταν από πολλαπλούς κλιματικούς παράγοντες και αναρωτιόμασταν από ποιους παράγοντες εξαρτιόταν.

6. Στην παράσταση έλεγαν δυνατά ότι το δηλητήριο προοριζόταν για κάποιους από την βασιλική οικογένεια, αλλά δεν αποκαλύφθηκε μέχρι το τέλος για ποιους προοριζόταν.

7. Χθες το βράδυ οι γονείς του παιδιού τσακώθηκαν για κάποιον, αλλά δεν είχε συναίσθηση για ποιον τσακώθηκαν.

8. Πριν εξαφανιστεί, η μικρή Αννούλα επέμενε να παίξει με κάποιους από το σχολείο, ωστόσο η μητέρα της δε θυμόταν με ποιους επέμενε να παίξει.

9. Οι αρχές ενημέρωσαν ότι στη νεροποντή δραπέτευσαν άγρια ζώα από κάποιον ζωολογικό κήπο, αλλά δεν πληροφορηθήκαμε από ποιον κήπο δραπέτευσαν.

10. Τα παιδιά της χορωδίας έλεγαν ότι η διευθύντρια μοιάζει με έναν διάσημο ηθοποιό, και μου φαίνεται πως κατάλαβα με ποιον ηθοποιό μοιάζει.

11. Μετά το σκάνδαλο, η πανικοβλημένη πρόεδρος της εταιρείας σκόπευε να βάλει τα λεφτά σε διαφορετικό λογαριασμό, αλλά δεν μπορούσε να αποφασίσει σε ποιον λογαριασμό να τα βάλει.

12. Στο μάθημα της λογοτεχνίας οι μαθητές έπρεπε να πουν τη γνώμη τους για έναν συγκεκριμένο χαρακτήρα, αλλά ως συνήθως η Βάσω δεν είχε ιδέα για ποιον χαρακτήρα έπρεπε να πουν τη γνώμη τους.

13. Σε περίπτωση που θελήσει η μικρότερη κόρη να σπουδάσει έξω, η οικογένεια θα αναγκαστεί να δανειστεί από κάποιους λεφτά, απλώς δεν ξέρει ακόμα από ποιους θα δανειστεί λεφτά.

14. Η Μαρία εξοργίστηκε όταν της είπαν συμφοιτητές πως έδωσαν τις σημειώσεις της σε κάποιον, αλλά δήθεν δεν μπορούσαν να τη διαφωτίσουν σε ποιον τις έδωσαν.

15. Η ανθοπώλης είχε ως κρυφό στόχο της να διαφοροποιηθεί οπωσδήποτε από κάποιον γνωστό της, αλλά ποτέ δεν αποκάλυψε παραέξω από ποιον είχε στόχο να διαφοροποιηθεί.

16. Ο Οδυσσέας ήξερε ότι η Πηνελόπη πήγαινε κρυφά τα βράδια σε κάποιον μνηστήρα, αλλά έπρεπε κάπως να πληροφορηθεί σε ποιον μνηστήρα πήγαινε.

17. Το Σάββατο στη λαϊκή η μανάβισσα αστειευόταν με κάποιους γνωστούς της, αλλά ο άνδρας της δεν ήξερε με ποιους αστειευτόταν.

18. Το πρωί ο διευθυντής της εταιρείας απευθυνόταν εξαγριωμένος σε κάποιους στο τηλέφωνο, αλλά δεν μπορούσε κανείς μας να καταλάβει σε ποιους απευθυνόταν.

19. Ακούσαμε ότι για να εξασφαλίσει τη θέση του, ο προπονητής ποδοσφαίρου επωφελήθηκε ιδιαίτερα από κάποιους γνωστούς του στο χώρο, αλλά δεν ήταν ξεκάθαρο από ποιους επωφελήθηκε.

20. Πολύ συχνά η δασκάλα μας παραπονιόταν στο διευθυντή για κάποιους από την τάξη, αλλά δεν καταφέραμε ποτέ να μάθουμε για ποιους παραπονιόταν.

21. Η μητέρα μου αργότερα μου είπε ότι η συμπεριφορά μου κακοφάνηκε σε κάποιους από τους επισκέπτες, αλλά δεν ήθελε να με πληροφορήσει σε ποιους επισκέπτες κακοφάνηκε η συμπεριφορά μου.

22. Όταν κόπηκε δραματικά η σκηνή, η πρωταγωνίστρια της σειράς φάνηκε να στεναχωρήθηκε πολύ με κάποιον χαρακτήρα, αλλά δεν είδαμε με ποιον χαρακτήρα στεναχωρήθηκε.

23. Λόγω διαφωνιών, του χρόνου ο Μανόλης και η Αφροδίτη θέλουν να συγκατοικήσουν με διαφορετικούς φίλους, όμως αναρωτιούνται ακόμη με ποιους φίλους να συγκατοικήσουν.

24. Έξω από το αμφιθέατρο σήμερα η Δανάη φαινόταν να γελάει συνεχώς με κάποιον εκεί κοντά, ωστόσο δίσταζε να πει στους υπόλοιπους με ποιον ακριβώς γελούσε.

25. Λόγω των οφειλών του, ο τζογαδόρος στο προπατζίδικο αγωνιούσε για αρκετούς αγώνες, ωστόσο δεν είχε κανείς ιδέα για ποιους αγώνες συγκεκριμένα αγωνιούσε.

26. Την περσινή σεζόν λόγω τραυματισμού η νεαρή αθλήτρια απουσίαζε από πολλούς αγώνες, αλλά δεν μπορώ αυτή τη στιγμή να θυμηθώ από ποιους αγώνες ακριβώς απουσίαζε.

27. Στη ρομαντική κωμωδία η κόρη της πρωταγωνίστριας ήταν ερωτευμένη με κάποιον, αλλά δεν ήθελε με τίποτα να αποκαλυφθεί με ποιον ήταν ερωτευμένη.

28. Όταν είδε τα αποτελέσματα των μαθητών της η καθηγήτρια απογοητεύτηκε ιδιαίτερα με κάποιον, αλλά δεν άφησε να φανεί με ποιον απογοητεύτηκε.

29. Καθόλη τη διάρκεια της ζωής της η ηλικιωμένη καλόγρια προσευχόταν για κάποιον συγκεκριμένο, αλλά ποτέ δε μας αποκάλυψε για ποιον προσευχόταν.

30. Στο νέο της βιβλίο η συγγραφέας διαφωνεί έντονα με κάποιον από το χώρο της τέχνης, αλλά δεν έχω διαβάσει ακόμη με ποιον διαφωνεί.

31. Χθες στην πλατεία η ενοχλητική γειτόνισσα καυχιόταν σε όλους για έναν από τους γιους της, αλλά δεν έκατσα να ακούσω για ποιον γιο συγκεκριμένα καυχιόταν σε όλους.

32. Στις ειδήσεις έλεγαν πως η τηλεπαρουσιάστρια είχε θεραπευτεί από κάποιον άσχημο ιό, αλλά δε θυμάμαι να ανέφεραν από ποιον ιό είχε θεραπευτεί.

33. Η μητέρα περηφανευόταν κρυφά για έναν μόνο από τους γιους της, αλλά ποτέ δεν θα παραδεχόταν για ποιον γιο περηφανευόταν κρυφά.

34. Σε κάποια φάση στο σινεμά η κοπέλα είχε συγκινηθεί με κάποιον δακρύβρεχτο χωρισμό, αλλά η φίλη της δεν είχε ιδέα με ποιον χωρισμό είχε συγκινηθεί.

35. Μετά την έντονη διαφωνία, η γυναίκα απομακρύνθηκε σταδιακά από κάποιον από τη δουλειά της, αλλά δεν ήθελε να γίνει φανερό στους υπόλοιπους από ποιον απομακρύνθηκε σταδιακά.

36. Ο μαθηματικός είπε στη βοηθό του ότι ήταν βαθιά συγκεντρωμένος σε κάποιους σημαντικούς υπολογισμούς, αλλά δεν την ενημέρωσε σε ποιους υπολογισμούς ήταν συγκεντρωμένος.

37. Η μικρή αδερφή της Ματίνας παραδέχτηκε ότι ξετρελάθηκε με κάποιους παλιούς δίσκους μουσικής, αλλά ντρεπόταν να μας πει με ποιους δίσκους ξετρελάθηκε.

38. Μετά την παράσταση η πρώτη χορεύτρια υποκλίθηκε σε κάποιους στο πίσω μέρος της σκηνής, αλλά δεν μπορούσε κανείς να δει σε ποιους υποκλίθηκε.

39. Κάπου στη μέση του μαθήματος, η γκρινιάρα έφηβη διαμαρτυρήθηκε για κάποιον από το διπλανό θρανίο, αλλά η καθηγήτρια δεν άκουσε για ποιον διαμαρτυρήθηκε.

40. Η εκπαιδευόμενη ηθοποιός είχε πολλούς ρόλους στην ερχόμενη παράσταση και ήθελε να παραιτηθεί από κάποιον, αλλά δεν είχε αποφασίσει ακόμα από ποιον ήθελε να παραιτηθεί.

41. Μέσα στη σιωπή το μωρό μάλλον θα ξαφνιάστηκε με κάποιον θόρυβο, αλλά δεν είχαμε συναίσθηση με ποιον θόρυβο ξαφνιάστηκε.

42. Την ώρα των προφορικών, η καθηγήτρια Γαλλικών αγχώθηκε ιδιαίτερα για κάποιους από το τμήμα της, ωστόσο δεν παραδέχτηκε για ποιους αγχώθηκε ιδιαίτερα.

43. Στην εταιρική συνάντηση σχετικά με τη συμπεριφορά των υπαλλήλων, η διευθύντρια έκανε εκτενή αναφορά σε κάποιους, αλλά δεν πρόσεξα σε ποιους έκανε αναφορά.

44. Αργά το βράδυ η δεκτριάχρονη μιλούσε κρυφά σε κάποιον στο τηλέφωνο, και η μητέρα της είχε βαλθεί να μάθει σε ποιον μιλούσε.

45. Μετά την πτώση του χρηματιστηρίου η εταιρεία απολογήθηκε σε κάποιους από τους μετόχους της, αλλά ο διευθυντής δεν ήθελε να μας ανακοινώσει σε ποιους μετόχους απολογήθηκε η εταιρεία.

46. Η δικηγόρος της γυναίκας δεν ήθελε να συμβιβαστεί με ορισμένους όρους του διαζυγίου, και εύκολα καταλάβαινε κανείς με ποιους όρους δεν ήθελε να συμβιβαστεί.

47. Στην έκθεση τέχνης που επισκεφθήκαμε μια μεσήλικη κυρία διαμαρτυρήθηκε για κάποιον χυδαίο πίνακα, αλλά κανένας μας δεν ήταν σίγουρος για ποιον πίνακα διαμαρτυρήθηκε.

48. Παρά τις προτροπές μας, ξέρω ότι η αδερφή μου χθες μιλούσε σε κάποιους στο μπαρ, αλλά δεν είμαι σίγουρη ακριβώς σε ποιους μιλούσε.

49. Όταν γύρισε από την Ταϋλάνδη, η ταξιδιώτισσα αναγκάστηκε να μπει σε καραντίνα επειδή είχε εκτεθεί σε κάποιον ιό, αλλά δεν την ενημέρωσαν σε ποιον ιό ακριβώς είχε εκτεθεί.

50. Η μικρή εξερευνήτρια ήθελε να εξομολογηθεί το μεγάλο της μυστικό σε κάποιον φίλο της, αλλά δεν είχε αποφασίσει ακόμη σε ποιον φίλο ήθελε να το εξομολογηθεί.

51. Προτού προβεί σε επίσημο παράπονο, η γυμνάστρια ήθελε πρώτα να σιγουρευτεί για κάποιους από τους ισχυρισμούς εις βάρους της, αλλά δεν ήταν ξεκάθαρο για ποιους ισχυρισμούς ακριβώς ήθελε πρώτα να σιγουρευτεί.

52. Η πιανίστρια συνειδητοποίησε δυστυχώς ότι είχε αποτύχει σε κάποιους από τους στόχους της, αλλά δεν μας αποκάλυψε σε ποιους στόχους συνειδητοποίησε ότι είχε αποτύχει.

53. Μετά τη συμπεριφορά της στο γήπεδο η μπασκετμπολίστρια αποκλείστηκε από αρκετούς επερχόμενους αγώνες, αλλά δε θυμάμαι να σου πω από ποιους αγώνες αποκλείστηκε.

54. Για το επόμενο θύμα της, η μονομάχος εστίασε την προσοχή της σε κάποιον άλλον, όμως δεν ήταν εμφανές ακόμη σε ποιον εστίασε την προσοχή της.

55. Η μάνατζερ της πρωταθλήτριας ομάδας ποδοσφαίρου αποφάσισε κρυφά να επενδύσει πολλά λεφτά σε κάποιον, αλλά δεν ήθελε να ανακοινωθεί ακόμη σε ποιον αποφάσισε να επενδύσει.

56. Την ημέρα των βαθμών η φοβιτσιάρα δασκάλα κρυβόταν από κάποιους ενοχλητικούς γονιούς, και όλοι είχαν καταλάβει από ποιους γονιούς κρυβόταν.

57. Για την πρώτη της ανάθεση, η νέα δημοσιογράφος έπρεπε να πάρει συνέντευξη από κάποιον, αλλά δεν την είχαν ενημερώσει ακόμη από ποιον έπρεπε να πάρει συνέντευξη.

58. Στην αρχή της εξεταστικής, η φοιτήτρια είχε δανειστεί ένα στυλό από κάποιον, αλλά στο μεταξύ ξέχασε δυστυχώς από ποιον το δανείστηκε.

59. Η κόρη της Εύας της είπε ότι χαιρόταν πολύ που διέφερε τουλάχιστον από κάποιους στο τμήμα της, αλλά δεν της ανέφερε από ποιους συγκεκριμένα χαιρόταν που διέφερε.

60. Μπαίνοντας στο βιβλιοπωλείο η πελάτισσα έδειξε ενδιαφέρον για έναν τόμο ποίησης, αλλά ο μαγαζάτορας δεν αντιλήφθηκε για ποιον τόμο ακριβώς έδειξε ενδιαφέρον.

61. Στο αεροπλάνο από την Ισπανία, η Μάρθα είχε εντυπωσιαστεί πολύ με κάποιους από το πλήρωμα καμπίνας, αλλά οι υπόλοιποι δεν καταλάβαμε με ποιους ακριβώς είχε εντυπωσιαστεί πολύ.

62. Πριν αρκετό καιρό η δημοφιλής μαζορέτα είχε τσακωθεί με κάποιους από την ομάδα της, ωστόσο όλοι έχουν ξεχάσει πλέον με ποιους είχε τσακωθεί.

63. Το φυλλάδιο οδηγιών του νέου φαρμάκου τονίζει πως ενδείκνυται για συγκεκριμένους ασθενείς μόνο, αλλά δεν ήταν ιδιαίτερα ξεκάθαρο για ποιους ασθενείς ενδείκνυται.

64. Η νέα νοσοκόμα πανικοβλήθηκε όταν έμαθε ότι ευθύνεται σήμερα για κάποιους στον χώρο της εντατικής, αλλά δεν μας ενημέρωσε για ποιους ευθύνεται.

B.1.1.2 Fillers

1. Η ξαδέρφη του χρειάστηκε να απολογηθεί επανειλημμένα για τη συμπεριφορά του νεαρού ενός και ο αδερφός της για του άλλου.

2. Η ντροπαλή της παρέας ενθουσιάστηκε όταν είπαν οι άλλοι ότι θέλανε να πάνε σινεμά, αλλά δε θυμόταν τι είπαν ότι προβαλόταν εκείνη την ημέρα.

3. Η κόρη της πωλήτριας μιλούσε άσχημα τη μία μέρα για τον έναν πελάτη, και την άλλη μέρα για την άλλη, προς απογοήτευση της μητέρας της.

4. Η γρια περίμενε συχνά τους δύο γιους της την Κυριακή για μεσημεριανό και σκεφτόταν όλη την εβδομάδα τι να τους μαγειρέψει.

5. Η μητέρα του είναι πολύ περήφανη για τον Αριστείδη, και φαίνεται από το πόση ώρα κάθεται και καυχιέται στις γειτόνισσες καθημερινά.

6. Οι δύο κοπέλες είχαν τσακωθεί άσχημα για τον Αντώνη, αλλά δε θυμάμαι πια εδώ και πόσον καιρό είχαν ψυχραθεί.

7. Πριν το γάμο η νύφη δοκίμασε πολλά χτενίσματα, διαλέγοντας ένα, αλλά δεν κατάλαβα πόση ώρα θα την έπαιρνε να χτενιστεί την ημέρα του γάμου.

8. Το κοριτσάκι χάιδευε για πολλή ώρα του ενός αδέσποτου κουταβιού, ενώ ο αδερφός της έκανε μπάνιο το άλλο.

9. Η μαθήτρια έμοιαζε ολοφάνερα στον πατέρα της και η μητέρα της αναρωτιόταν πώς γινόταν να μην της μοιάζει καθόλου.

10. Στη συνέντευξη τύπου, η διευθύντρια αναφέρθηκε εκτενώς σε κάποια καλές πρακτικές της εταιρείας, ενώ αγνόησε επίτηδες κάποιες κακές.

11. Η δικηγόρος μάζεψε τα χαρτιά της και ο δικαστής τα δικά του προτού φύγει μαζί να πάνε να φάνε στο κυριλέ εστιατόριο της περιοχής.

12. Η γειτόνισσα δανείστηκε λεφτά από τη φίλη της για να στο θέατρο δώσει γιατί ξέχασε το πορτοφόλι της.

13. Η αρχιτέκτονας είχε σχεδιάσει ήδη τα δύο μεγάλα παράθυρα του σπιτιού, αλλά ήθελε τη γνώμη των πελατών για τα υπόλοιπα.

14. Η κόρη της Ελένης πήρε δώρο στο αγόρι της και ήθελε να πάρει και στην αδερφή του και έτσι έκανε κύκλους στα μαγαζιά.

15. Η πωλήτρια πήγαινε κρυφά κάποιες φορές στο θέατρο να δει τους αγαπημένους της παραστάσεις, αλλά δεν θυμάμαι ποια ήταν ακριβώς.

16. Η αδερφή της Άννας φώναζε πολύ στο τηλέφωνο όταν πήγαμε για καφέ, αλλά δεν ήθελε να μου πει γιατί ήταν τόσο θυμωμένη.

17. Η νοσοκόμα απευθυνόταν θυμωμένα στο Γιάννη, όμως δεν μπορούσα να δω ποιος άλλος είχε στοχοποιηθεί από τη νοσοκόμα.

18. Η συγγραφέας κατάφερε να εκδώσει η ιστορία αγάπης της στον έναν μεγάλο εκδοτικό οίκο που ήθελε και τα ποιήματά της στον άλλον.

19. Η ξαδέρφη του παραπονιόταν συνεχώς για το Μάνο και ο αδερφός του το ίδιο και με το παραπάνω.

20. Η επιμελήτρια έδιωξε κακήν κακώς έναν μαθητή από τη μεγάλη αίθοσα, αλλά δεν είδε τους άλλους δύο που μπήκαν κρυφά.

21. Η δασκάλα χάρηκε με την πρόοδο ενός μαθητή, αλλά στεναχωρήθηκε με τη στασιμότητα άλλους πέντε.

22. Οι δύο φοιτήτριες συγκατοικούσαν ήδη με τη μία αδερφή τους και δεν ήθελαν με τίποτα να συγκατοικήσουν και με την άλλη.

23. Η ιδιοκτήτρια του καταστήματος παράγγειλε τριάντα κουτία καφέ, αλλά ξέχασε δέκα τα ζάχαρης που ήθελε.

24. Η καθηγήτρια ξένων γλωσσών αγχώθηκε οι εξετάσεις των μαθητών της πώς να τα πήγαν τελικά.

25. Η νοικοκυρά έδειχνε στα δύο μικρά παιδιά της πώς να στρώθηκε το τραπέζι για να αρχίσουν να βοηθάνε.

26. Η τουρίστρια ερωτεύτηκε το πανέμορφο κτίριο στο κέντρο της πόλης και έψαχνε ποιος θα μπορούσε να την ξεναγήσει.

27. Καθώς σχεδίαζε η διακοσμήτρια το νέο σαλόνι του ζευγαριού, αναρωτιόταν τι χρώμα θα τους άρεσε καλύτερα στον έναν τοίχο.

28. Η παρουσιάστρια πήρε συνέντευξη σήμερα από τον νεαρό υπουργό, αλλά δεν άκουσα γιατί εκνευρίστηκε τόσο μαζί του.

29. Η νταντά ευθύνεται τα μεσημέρια για τον Κωστάκη και η μεγαλύτερη αδερφή του για τον οικογενειακό γάτο που όλο κάνει ζημιές.

30. Η αρραβωνιαστικιά του ανυπομονούσε προφανώς να δει το Νικόλα, και σκεφτόταν όλο το πρωί τι φαγητό να του φτιάξει όταν έρθει.

31. Η γιαγιά συμπαραστεκόταν πάντα στοϊκά στο νεαρό εγγονό της και ο παππούς του το ίδιο, αλλά χωρίς να το δείχνει τόσο.

32. Η θεία έπρεπε να εξηγήσει την ιστορία πάλι στον Κίμωνα, αλλά δεν μπορώ να θυμηθώ τώρα ποιος άλλος ενδιαφερόταν να τη μάθει.

33. Η κόκα κόλα σερβίρεται ακόμα παντού παρόλο που πολλές έρευνες έχουν δείξει πόσο ανθυγιεινή είναι ειδικά για τα μικρά παιδιά.

34. Η λογίστρια εκνευρίστηκε αμέσως με τον Παναγιώτη, και θα ήθελα να μάθω γιατί ακριβώς αρπάχτηκαν οι δυο τους ξαφνικά.

35. Η αστυνόμος επιτέθηκε με θάρρος στον έναν κλέφτη και ο σκύλος του έριξε κάτω τον άλλον και έτσι η αστυνομία συνέλαβε και τους δύο.

36. Όταν το αστείο παρατράβηξε, η κοπέλα του αναγκάστκε να εξηγηθεί εντέλει στο Μάρκο και η Μαρία στο νεαρό της γιατί είχαν νευριάσει και οι δύο.

37. Η φεμινίστρια τρελαινόταν κρυφά για το Θανάση, αλλά δεν μπορούσε να καταλάβει πώς μαθεύτηκε αυτό παραέξω.

38. Η αδερφή του Κωνσταντίνου ανυπομονούσε να μεγαλώσει ώστε να επιτρέπεται να βγαίνει αργά με τους φίλους της όπως αυτός.

39. Η προπονήτρια διαφώνησε έντονα με έναν παίχτη και ο μάνατζερ δυστυχώς το ίδιο προκαλώντας μια άσχημη ατμόσφαιρα στην προπόνηση.

40. Η νονά προσευχόταν κάθε βράδυ για το Μίλτο και η μαμά του για το Μάρκο, μιας και ήταν και οι δύο πολύ θρησκευόμενες.

41. Η επιτροπή ανέκρινε τη Μαρίκα σχετικά με το ξέπλυμα μαύρου χρήματος και το Μάριο για τις πλαστές επιταγές, αλλά δεν κατάφεραν άκρη.

42. Η βιβλιοπώλης συχνά διάλεξε ένα νέο βιβλίο για τη μία φίλη της και ένα παλιότερο για την άλλη σαν δώρο Χριστουγέννων.

43. Η νονά της Μαρτίνας τελικά υπέκυψε και αναγκάστηκε να της πάρει την πιο ακριβή λαμπάδα του μαγαζιού.

44. Η μητέρα του Τοτού εκμεταλλεύτηκε την αγάπη του για τις σοκολάτες ώστε τον κάνει να διαβάσει να.

45. Ο Αχιλλέας λαβώθηκε θανάσιμα στη φτέρνα του τα βέλη του Πάρι από, και έτσι βγήκε η έκφραση «Αχίλλειος πτέρνα».

46. Ο υπουργός στη συνέντευξη τύπου είπε ότι αν και θα καταργηθούν κάποιες περικοπές, θα εισαχθούν συμπληρωματικούς σε άλλοι τομείς.

47. Η Μαρία παραπονιόταν ότι πολλά μέτρα έχουν παρθεί τα τελευταία χρόνια για την ανακύκλωση, αλλά ελάχιστα υλοποιούνται.

48. Η Χριστίνα αγχωνόταν πού είχε βάλει την τσάντα της και πού το πορτοφόλι μιας και βιαζόταν για τη συνάντησή της.

49. Είχαν γίνει πια τόσες απεργίες ώστε ο κόσμος ελάχιστα συγκινούνταν πέρα την ταλαιπωρία από που τους προκαλούσε.

50. Ο φαρμακοποιός λυπήθηκε τη φτωχή ζητιάνα και της έδωσε τα φάρμακα που χρειαζόταν να την αφήσει να πληρώσει δίχως.

51. Ο εκτυπωτής είχε ξεμείνει πάλι από μελάνι, αλλά ο Ανέστης δεν είχε όρεξη να πάει να αγοράσει νέα.

52. Η αλεπού κυνηγούσε για πολλή ώρα του κουνελιού ώσπου να καταφέρει να το στριμώξει και να το πιάσει.

53. Ο ζωγράφος εμπνεύστηκε από την ηρεμία της γάτας του που γουργούριζε καθώς κοίταζε μονίμως από το παράθυρο.

54. Η Μάρθα κοίταζε τον εαυτό της στο καθρέφτη και έκανε γκριμάτσες, ενώ ο αδερφός της γελούσε.

55. Μέσα στη σιωπή, το ρολόι ηχούσε δυνατά και τραβούσε την προσοχή του μωρού ανήσυχου .

56. Όσες φορές κι αν μάζευε η Νικολέτα τα παιχνίδια από το πάτωμα του σαλονιού, πάντα κατάφερναν να ξαναβρεθούν άμεσα εκεί.

57. Η απόφαση πάρθηκε τελικά το συμβούλιο από να σταματήσουν την εξαγωγή προϊόντων προς τις συγκεκριμένες χώρες.

58. Ο νεαρός μαθητής αναρωτιόταν πώς μπορούσαν να αντέχουν οι μαθηματικοί τόσες μεταβλητές κάθε μέρα να παθαίνουν εγκεφαλικό χωρίς.

59. Η ντίβα αναρωτιόταν πότε θα έφταναν επιτέλους τα νέα καλλυντικά που παρήγγειλε για να πετάξει τα παλιά.

60. Ο σκηνοθέτης χτυπιόταν όλο το απόγευμα τα καμώματα των ηθοποιών με που δε συγκεντρώνονταν με τίποτα στο έργο.

61. Το παιδί του γείτονα κυνηγήθηκε επανειλημμένα από το σκύλο μας με αποτέλεσμα να αποκτήσει χρόνιο ψυχολογικό πρόβλημα.

62. Ούτε που κατάλαβα πώς κλάπηκε την τσάντα μου το διαβατήριό μου από καθώς και το ρολόι μου.

63. Το μεγαλύτερο μέρος της γης καλύπτεται από νερό, κι αυτό ακόμα δεν έχει εξερευνηθεί εξ ολοκλήρου.

64. Όχι μόνο πέτρες, αλλά εκατοντάδες αντικείμενα άρχισαν να εκτοξεύονται από οι καταπέλτες προς το τέλος της μάχης.

65. Ο ηθοποιός είδε το αμάξι της Άννας να πηγαίνει στο ξενοδοχείο και ο σκηνοθέτης τη μηχανή του Γιώργου.

66. Μελέτες δείχνουν πως οι εκπομπές διοξειδίου του άνθρακα στη χώρα μας έχουν τριπλασιαστεί την τελευταία δεκαετία.

67. Μόλις δημιουργήθηκε ο πρώτος ψηφιακός χάρτης τον αναπτυσσόμενο εγκέφαλο για, προκαλώντας μεγάλο ενθουσιασμό στην επιστημονική κοινότητα.

68. Σε μία παράξενη σειρά γεγονότων, το κουνέλι του διπλανού σπιτιού κατάφερε να δαγκώσει το σκύλο μας και εκείνος την ουρά του.

69. Στη γιορτινή σύναξη, ο Γρηγόρης έφερε σπανακόπιτα, η Ματίνα μια μηλόπιτα και ο σκύλος τους μία μικρή μπάλα.

70. Ο δημοσιογράφος φύτεψε μια ροδιά στον κήπο του, και η ζηλιάρα Βαγγελιώ το ίδιο, αλλά ποιους λόγους δεν μπορώ να καταλάβω.

71. Ο παρουσιαστής στις ειδήσεις εξήγησε πώς η αστυνομία προχώρησε στις συλλήψεις των τρομοκρατών και πώς των συνεργατών τους.

72. Παιδάκι το στο αεροπλάνο γκρίνιαζε γιατί η αδερφή του είχε νέο παιχνίδι ενώ το δικό του είχε χαλάσει στο αεροδρόμιο.

73. Σύμφωνα με τον ετήσιο έκθεση του σχολείου, ο Αναστάσης είχε τους πιο υψηλούς βαθμούς πέρυσι και η Μαρία φέτος.

74. Πριν από μερικά χρόνια είχε πιαστεί ένα κολοσσιαίο καλαμάρι μήκους πέντε μέτρων από ντόπιους ψαράδες έξω από τη Νέα Ζηλανδία.

75. Διερευνήθηκε πρόσφατα από επιστήμονες κορυφαίου ευρωπαϊκού πανεπιστημίου το πώς των σκύλων μπορούν να μιμηθούν τις κινήσεις άλλων σκύλων γύρω τους.

76. Πρόσφατα ένας φημισμένος προπονητής ποδοσφαίρου ενεπλάκη σε σκάνδαλο δωροδοκίας, αλλά δεν άκουσα αν αποκλείστηκε από το άθλημα ή όχι.

77. Στην οικογενειακή γιορτή οι γονείς μετά βίας τελείωσαν μία μεσαία τούρτα, ενώ τα παιδιά κατασπάραξαν πέντε.

78. Ο γείτονας είδε το Γρηγόρη να καθαρίζει το του αμάξι νέο και θυμήθηκε ότι έπρεπε κι αυτός το ίδιο.

79. Οι πλυμμήρες κατέστρεψαν τις σοδειές πολλών κατοίκων της περιοχής, ενώ προκάλεσαν ταυτόχρονα ζημιές στα σπίτια άλλων.

80. Στα βόρεια πολλά βουνά παραμένουν χιονισμένα καθόλη τη χρονιά, ενώ νότια έχουν χιόνι μόνο τις πιο κρύες χειμωνιάτικες μέρες.

81. Η κυβέρνηση στις τελευταίες συναντήσεις της σπρώχνει για αξιοποίηση της δημόσιας ακίνητης περιουσίας, αλλά δεν είναι σίγουρο ακόμα ποιες χώρες θα ήθελαν να επενδύσουν σε αυτήν.

82. Μια γνωστή εταιρεία ανήγγειλε τη δημιουργία νέων υποδημάτων που σε καθοδηγούν με δονήσεις στον προορισμό σου, και στη συνέντευξη τύπου κάποιοι δημοσιογράφοι παραξενεύτηκαν ενώ άλλοι ενθουσιάστηκαν.

83. Πρόσφατη έρευνα έδειξε ότι τα όνειρα είναι πολύ σημαντικά στην παγίωση των αναμνήσεων, και εάν διαταραχθούν υπάρχει περίπτωση να μη λειτουργεί καλά η μνήμη μας την επόμενη ημέρα.

84. Τα πλαϊνά τζάμια δεν προστατεύουν καλά από την υπεριώδη ακτινοβολία του ήλιου, οπότε οι επιστήμονες συνιστούν να φοράμε αντηλιακό και μακριά μανίκια στο αμάξι για προστασία.

85. Η ξαδέρφη του Γρηγόρη είπε ότι τα παιδιά που περπατάνε νωρίτερα έχουν πιο γερά κόκαλα, οπότε είχαμε περιέργεια να δούμε πότε θα περπατούσε η μικρότερη κόρη μας.

86. Η Αναστασία έλεγε ότι στην Κροατία έχουν πολλά έθιμα παρόμοια με τα ελληνικά για το Πάσχα, και μιλούσε για το τσουρέκι και τα κόκκινα αυγά που είδε εκεί.

87. Λόγω του ανθυγιεινού τρόπου ζωής της μοντέρνας κοινωνίας η παχυσαρκία αποτελεί πλέον πραγματική επιδημία, ενώ εκτιμάται ότι σύντομα θα πάσχουν περισσότεροι από τους μισούς Ευρωπαίους.

88. Η δασκάλα του Λουκά άργησε να φτάσει στο μάθημα λόγω ενός ατυχήματος στην Εθνική Οδό, αλλά ευτυχώς μαθεύτηκε σύντομα στις ειδήσεις ότι κανείς δεν είχε τραυματιστεί σε αυτό.

89. Η μικρή Ηλιάνα έμαθε στο σχολείο σήμερα για κάποια νέα έντομα που ανακαλύφθηκαν πρόσφατα, και δεν μπορούσε να σταματήσει να γελάει όταν άκουσε πώς τα ονόμασαν οι επιστήμονες.

90. Ο καθηγητής της φίλης μου εξηγούσε πως ο διαστημικός καιρός επηρεάζει τον γήινο καιρό, και συζητούσαμε πώς θα μελετούνταν αυτά τα δύο μαζί.

91. Ο γιατρός εξηγούσε στη μητέρα μου ότι υπάρχει καλή και κακή χοληστερόλη, όμως αυτή επιμένει να τρώει ανθυγιεινά και να καπνίζει προς μεγάλη μας απογοήτευση.

92. Το κλειδί για ένα πετυχημένο δώρο δεν είναι να σκέφτεσαι τον εορτάζοντα, αλλά να αγοράζεις αυτό που αρέσει σε σένα για να αποτυπώνεται σε αυτό η προσωπικότητά σου.

93. Η μητέρα του Γιωργάκη ήθελε να του πάρει ένα σκυλάκι ως έκπληξη για τη γιορτή του, αλλά δεν μπορούσε να αποφασίσει ποιο σκύλο να σώσει τη φιλοζωική οργάνωση από.

94. Η γάτα του γείτονα επιτέθηκε στη δική μας γάτα όταν την πρωτοπήραμε, όμως δε θα μπορούσε να φανταστεί κανείς ότι θα γίνονταν σύντομα οι καλύτεροι φίλοι.

95. Ο ταχυδρόμος της γειτονιάς παρατηρήσαμε πως κάνει πολύ γρήγορα τη δουλειά του, και αναρωτιόμασταν μήπως προσπερνούσε κάποια σπίτια για να τελειώσει γρήγορα πιο πολύ.

96. Η Ντορέτα αγχωνόταν και έτρεχε πάνω κάτω μέσα στο σπίτι βιαστικά πριν το ραντεβού της, αλλά δεν μπορούσε να βρει το ωραίο μπουφάν της.

97. Η κοπέλα κοίταζε μελαγχολικά το κινητό περιμένοντας μήπως τον πάρει τηλέφωνο ο αγαπημένος της, ωστόσο μετά από κάποια ώρα συνειδητοποίησε ότι περίμενε άδικα γιατί δεν είχε σήμα.

98. Οι καθοδηγημένες εξορμήσεις με μοτοσυκλέτες είναι μια νέα μορφή τουρισμού, αλλά δεν μπορώ να φανταστώ ότι κάποιος τους γνωστούς μας από θα δελεαζόταν από αυτό.

99. Η κυβέρνηση κάνει προσπάθεια κάθε να περιορίσει τις εγκληματικές δραστηριότητες στη χώρα μας, όμως μελέτες δείχνουν πως αντιθέτως έχουν πολλαπλασιαστεί κατά την τελευταία δεκαετία.

100. Ο ζωγράφος της παρέας πήγαινε κάθε πρωί βόλτα στο πάρκο για έμπνευση, αλλά σήμερα εμπνεύστηκε από τον γάτο που καθόταν το παράθυρό του έξω από.

101. Η αστυνομία ανέκρινε τους υπόπτους σχετικά οι ληστείες με και ο δικαστής αργότερα το ίδιο, αλλά κανένας τους δεν ομολόγησε τίποτα.

102. Ο αδερφός της Πολυξένης είχε ενθουσιαστούμε πολύ με τους χειμερινούς Ολυμπιακούς αγώνες στην τηλεοράση, όμως οι υπόλοιποι δε γνωρίζαμε κανέναν από τους αθλητές και έτσι βαρεθήκαμε γρήγορα.

103. Η πωλήτρια μας έλεγε των τιμών των προϊόντων ομορφιάς και τις προσφορές, όμως ακόμα δεν κατάλαβα γιατί η κρέμα χεριών πωλείται πολύ πιο ακριβά από την κρέμα σώματος.

104. Ο καθηγητής μας έλεγε ότι δημιουργήθηκε ο πρώτος ψηφιακός χάρτης για τον αναπτυσσόμενο εγκέφαλο, και προκλήθηκε μεγάλος ενθουσιασμός στην επιστημονική κοινότητα ως συνέπεια αυτό το μεγάλο κατόρθωμα.

105. Οι γονείς του κοριτσιού μάλωσαν έναν από τους συμμαθητές της όταν την έφεραν στο σχολείο, αλλά δεν ακούσαμε ποιους λόγους τον μάλωσαν γιατί βρισκόμαστα αρκετά μακριά.

106. Η Πηνελόπη εξοργίστηκε όταν της είπαν οι γονείς της ότι έδωσαν το βιβλίο της κάπου, αλλά δεν μπόρεσε να εξακριβώσει πού το έδωσαν παρόλο που προσπαθήσουν πολύ.

107. Στη γνωστή τηλεοπτική σειρά όλοι ήξεραν ότι κάποιος θα πέθαινε μέχρι το τέλος της σεζόν, αλλά έπρεπε να περιμένουμε να μάθουμε ποιος θα σκότωνε από ποιον χαρακτήρα.

108. Ο δημοσιογράφος είχε αγανακτήσει πια με τις ανοησίες της συμπαρουσιάστριάς του, αλλά δεν ήξερε πότε θα προσλαμβανόταν νέα παρουσιάστρια στο κανάλι ώστε να μπορέσει να ησυχάσουμε λίγο.

109. Ο υπαστυνόμος τη συνέντευξη τύπου σε είπε στους δημοσιογράφους ότι είχαν κάποιους υπόπτους, αλλά δεν αποκάλυψε ποιοι από αυτούς έπρεπε να διερευνηθούν περαιτέρω για αδιάσειστα ενοχοποιητικά στοιχεία.

110. Οι θεατές ανυπομονούσαν να αποκαλυφθεί ο ένοχος στην τελευταία σκηνή της σειράς, αλλά τελικά δεν αποκαλύφθηκε από ποιον δολοφονήθηκε οι αδερφές της πρωταγωνίστριας προς αγανάκτηση όλων.

111. Οι αρχαιολόγοι είχαν κάνει μεγάλες προόδους στις ανασκαφές της αρχαίας πόλης, ωστόσο παραδέχτηκαν ότι μπορεί ποτέ να μη μάθουμε πώς ακριβώς καταστράφηκε αυτός.

112. Οι εργαζόμενοι νοσοκομείου του είχαν κάνει ήδη πολλές διαμαρτυρίες και απεργίες, ωστόσο κανείς δεν ήξερε ακόμα πότε θα πληρώνονταν ή αν θα έμενε καν ανοιχτό το νοσοκομείο.

113. Τα λαμπραντόρ λέγεται ότι είναι η πιο υπάκουες ράτσες σκύλου, ωστόσο εμείς προτιμήσαμε να πάρουμε λυκόσκυλο διότι θέλαμε περισσότερο έναν φύλακα για το σπίτι.

114. Η γιαγιά καιρό να υιοθετήσει μια γάτα και τελικά το έκανε, αλλά κάθε καλοκαίρι το μετανιώνει διότι μαδάει τόσο πολύ που συνέχεια καθάρισμα το σπίτι θέλει.

115. Ένα ιδιαίτερα ηρεμιστικό χόμπι λένε ότι είναι το η ζωγραφική, και μετά από πολύ κόπο καταφέραμε να πείσουμε τη μαμά να αρχίσει το για να ηρεμεί το μεσημέρι.

116. Η κοπέλα από μικρή διάβαζε κάθε βιβλίο που πέρναγε από τα χέρια της, και έτσι δε θα μπορούσα να σου πω τι διαβάζει τώρα με απόλυτη σίγουρος.

117. Η Λουίζα φοβάται πολύ τα φίδια, και οι γονείς της την πήγανε σε ένα ζωολογικό κήπο να τα δει από κοντά ώστε να ξεπεράσει το φόβο της.

118. Κάθε παιδί λένε ότι ονειρεύεται να γίνει αστροναύτης, αστυνόμος ή πυροσβέστης όταν μεγαλώσει, όμως αναρωτιέμαι πόσα παιδιά ονειρεύονται να γίνουν μαγείρων ή βιβλιοθηκαρίων.

119. Η κηπουρός του ξενοδοχείου χαιρόταν πολύ να φροντίζει τα τριαντάφυλλα στην μπροστινή είσοδο, αλλά δεν ήταν ξεκάθαρο εάν είχε και δικό της προσωπικός κήπος στο σπίτι.

120. Ο μάγειρας του εστιατορίου έφτιαξε μία πεντανόστιμη πίτα για κυρίως πιάτο, και συζητούσαμε το πόσο ζηλεύαμε την οικογένειά του που έτρωγε τόσο ωραίο φαγητό καθημερινά.

121. Λένε ότι εάν βάλεις αλάτι στο νερό θα αργήσει περισσότερο να έρθει σε βράση, αλλά η ξαδέρφη μου ακόμα δεν μπορεί να καταλάβει συμβαίνει αυτό γιατί.

122. Ο γαμπρός κοίταζε αραιά και πού την ώρα να μην αργήσει για, αλλά δεν κατάλαβε ότι το ρολόι του είχε σταματήσει πριν αρκετή ώρα και τελικά καθυστέρησε.

123. Η παιδίατρος είχε συμβουλέψει τη μαμά του Βασίλη να σταματήσει ο μικρός των πατατακίων, αλλά κανείς δεν κατάλαβε ότι εκείνος τα αγόραζε κρυφά από το περίπτερο.

124. Η νύφη του Ανέστη λάτρευε το τρέξιμο και πήγαινε έξω στις έξι κάθε πρωί, αλλά κανείς δεν είχε αντιληφθεί ακόμα πού ακριβώς πήγαινε για να τρέξω.

125. Ο νταής της γειτονιάς έτρεξε κλαίγοντας στο σπίτι του χθες το βράδυ, αλλά ντρεπόταν να πει σε κανέναν ότι χτυπήθηκε το μικρότερο κορίτσι της γειτονιάς από.

126. Ο Ραφαήλ είπε σε όλους ότι τραυματίστηκε σε μια σκληρή μάχη με άγριους εχθρούς, αλλά της αλήθειας είναι ότι γλίστρησε στο μπάνιο και έπεσε.

127. Οι γριούλες της γειτονιάς μαζεύονταν κάθε Κυριακή για να παίξουν χαρτιά, αλλά δεν ήταν ξεκάθαρο ποια από όλες εξαπατώνταν περισσότερο τις άλλες από.

128. Το κτίριο στη γωνία λίγο πιο κάτω από το σπίτι μας είναι από τους παλαιότερους στην περιοχή, αλλά είναι μυστήριο από ποιον χτίστηκε ακριβώς και για ποιο λόγο.

B.1.1.3 Plausibility Stimuli

Original Ferreira (2003) stimuli with Greek translation.

The last 8 stimuli are original sentences constructed following Ferreira (2003). One of the biased sentences (number 22) and two of the completely implausible sentences (numbers 5 & 22) were changed as they did not work as well in Greek. The short stimuli are then followed by an extended, more complex version (half passive, half active) to match the length and complexity of the Greek stimuli used in Experiment 1.

Biased reversible sentences (Short)

1. i. The man bit the dog.
   ii. Ο άνδρας δάγκωσε το σκύλο.
2. i. The food ruined the cook.
   ii. Το φαγητό κατέστρεψε το μάγειρα.
3. i. The worm ate the bird.

ii. Το σκουλήκι έφαγε τον παπαγάλο.

4. i. The mouse chased the cat.

   ii. Το ποντίκι κυνήγησε τη γάτα.

5. i. The villager protected the soldier.

   ii. Ο χωριάτης προστάτεψε τον στρατιώτη.

6. i. The doctor sued the lawyer.

   ii. Ο γιατρός άσκησε μήνυση εναντίον του δικηγόρου.

7. i. The student quizzed the teacher.

   ii. Ο μαθητής εξέτασε το δάσκαλο.

8. i. The thief pursued the cop.

   ii. Ο κλέφτης κυνήγησε τον αστυνόμο.

9. i. The man served the waitress.

   ii. Ο άνδρας σέρβιρε τη σερβιτόρα.

10. i. The cat fed the owner.

    ii. Η γάτα τάισε τον ιδιοκτήτη.

11. i. The suspect investigated the detective.

    ii. Ο ύποπτος ανέκρινε τον υπαστυνόμο.

12. i. The patient treated the doctor.

    ii. Ο ασθενής θεράπευσε το γιατρό.

13. i. The voter deceived the politician.

    ii. Ο ψηφοφόρος παραπλάνησε τον πολιτικό.

14. i. The mosquito killed the hiker.

    ii. Το κουνούπι σκότωσε τον ορειβάτη.

15. i. The rider threw the horse.

    ii. Ο αναβάτης πέταξε κάτω το άλογο.

16. i. The ball hit the golfer.

    ii. Η μπάλα χτύπησε τον παίχτη του γκολφ.

17. i. The deer shot the hunter.

ii. Το ελάφι πυροβόλησε το θηρευτή.

18. i. The fly ate the frog.
    ii. Η μύγα έφαγε τον βάτραχο.
19. i. The boy scared the ghost.
    ii. Ο νεαρός τρόμαξε το φάντασμα.
20. i. The jockey kicked the horse.
    ii. Ο αναβάτης κλώτσησε το άλογο.
21. i. The fish caught the angler.
    ii. Το ψάρι έπιασε τον ψαρά.
22. i. The bull dodged the matador.
    ii. Ο κυνηγός ξέφυγε από την πάπια.
23. i. The citizen arrested the officer.
    ii. Ο πολίτης συνέλαβε τον αστυνομικό.
24. i. The dragon slayed the prince.
    ii. Ο δράκος σκότωσε τον ιππότη.
25. Το χωριό λεηλάτησε τους πειρατές.
26. Η γη πέρασε τον κομήτη.
27. Η γυναίκα γραντζούνησε τη γάτα.
28. Το κουνέλι χάιδεψε την κοπέλα.
29. Το κοριτσάκι κούρεψε τον κομμωτή.
30. Ο σκύλος πήγε βόλτα το παιδί.
31. Οι ηθοποιοί χειροκρότησαν το θεατή.
32. Ο ποδοσφαιριστής σφύριξε στο διαιτητή.

Completely Implausible/Non-reversible sentences (Short)

1. i. The apron wore the chef.
   ii. Η ποδιά φόρεσε το μάγειρα.

2. i. The corn planted the farmer.
   ii. Η ελιά φύτεψε τον αγρότη.
3. i. The cheese ate the mouse.
   ii. Το τυρί έφαγε τον αρουραίο.
4. i. The bone buried the dog.
   ii. Το κόκαλο έθαψε το σκύλο.
5. i. The paper reviewed the editor.
   ii. Το άρθρο έγραψε το δημοσιογράφο.
6. i. The joke told the comic.
   ii. Το ανέκδοτο διηγήθηκε τον κωμικό.
7. i. The drain fixed the plumber.
   ii. Ο σωλήνας έφτιαξε τον υδραυλικό.
8. i. The race won the runner.
   ii. Ο αγώνας κέρδισε τον αθλητή.
9. i. The cud chewed the cow.
   ii. Το χορτάρι μάσησε την αγελάδα.
10. i. The music played the DJ.
    ii. Η μουσική έβαλε τον ντιτζέι.
11. i. The letter typed the secretary.
    ii. Η επιστολή έγραψε τη γραμματέα.
12. i. The picture painted the artist.
    ii. Το έργο τέχνης έβαψε/ζωγράφισε τον ζωγράφο.
13. i. The food chewed the termite.
    ii. Το δέντρο έφαγε τον τερμίτη.
14. i. The egg laid the chicken.
    ii. Το αυγό γέννησε την κότα.
15. i. The hill built the ant.
    ii. Η φωλιά έχτισε το μυρμήγκι.

16. i. The X-ray took the doctor.

    ii. Η ακτινογραφία έβγαλε το γιατρό.

17. i. The treasure buried the pirate.

    ii. Ο θησαυρός έθαψε τον πειρατή.

18. i. The shot gave the nurse.

    ii. Η ένεση έκανε τη νοσοκόμα.

19. i. The tooth pulled the dentist.

    ii. Το δόντι τράβηξε το γιατρό.

20. i. The tree gnawed the beaver.

    ii. Το δέντρο έφαγε τον κάστορα.

21. i. The wagon pulled the kid.

    ii. Ο φάκελος έστειλε τον επιχειρηματία.

22. i. The dirt pushed the bulldozer.

    ii. Το χώμα μάζεψε την μπουλντόζα.

23. i. The skirt hemmed the tailor.

    ii. Η φούστα έραψε τη μοδίστρα.

24. i. The plane flew the pilot.

    ii. Το αεροπλάνο πέταξε τον πιλότο.

25. Το ακόντιο πέταξε τον πρωταθλητή.

26. Η ανατολή του ήλιου φωτογράφισε τον νεαρό ρομαντικό.

27. Ο δρόμος ασφαλτόστρωσε το φορτηγατζή.

28. Η στέγη επισκεύασε το χτίστη.

29. Ο καθρέφτης έσπασε τη γυναίκα.

30. Το Χριστουγεννιάτικο δέντρο στόλισε τους ανθρώπους.

31. Η σούπα μαγείρεψε τον σεφ.

32. Η τάρτα λεμονιού δάγκωσε τον άνδρα.

Biased reversible sentences (Long)

1. Οι γείτονες ήθελαν να βάλει φίμωτρο ο άνδρας γιατί είχε δαγκώσει το πολυαγαπημένο σκύλο τους.
2. Ο μάγειρας ήταν στεναχωρημένος σε όλη τη διάρκεια του ραντεβού του γιατί νωρίτερα είχε καταστραφεί από το φαγητό του.
3. Η φώκια παρατηρούσε από μακριά το ευκίνητο σκουλήκι που έφαγε τον παπαγάλο για πρωινό.
4. Η γάτα κυνηγήθηκε από το χοντρό ποντίκι και ανέβηκε στο Χριστουγεννιάτικο δέντρο για να ξεφύγει.
5. Ο στρατιώτης προστατεύτηκε με θάρρος από το χωριάτη ενώ έτρεχε να σωθεί από τα βέλη των εχθρών.
6. Ο γιατρός άσκησε μήνυση εναντίον του δικηγόρου γιατί θεωρούσε πως δεν τον είχε αντιπροσωπεύσει στο έπακρο των ικανοτήτων του.
7. Ο δάσκαλος εξετάστηκε από το μαθητή μέσω δίωρου διαγωνίσματος που του έθεσε δίχως προειδοποίηση.
8. Ο κλέφτης κυνήγησε τον αστυνόμο μέσα από τους δρόμους του Λονδίνου, αλλά κατάφερε να του ξεφύγει προς μεγάλη του απογοήτευση.
9. Όταν ήταν έτοιμη, ο άνδρας πήρε την παραγγελία και σέρβιρε τη σερβιτόρα που καθόταν στο αγαπημένο της τραπέζι.
10. Όλη την ημέρα ο ιδιοκτήτης ήταν γεμάτος ανυπομονησία μέχρι να τον ταίσει ψαροκροκέτες η γάτα του.
11. Κατά τη διάρκεια της ανάκρισης ο υπαστυνόμος ανακρίθηκε από τον ύποπτο για τις πρόσφατες παράξενες κινήσεις του.
12. Ο γνωστός γιατρός θεραπεύτηκε από τον καραφλό ασθενή που είχε μόλις μπει στην εντατική έπειτα από τροχαίο.
13. Ο πολιτικός παραπλανήθηκε από τον ψηφοφόρο που του είχε υποσχεθεί μεγάλες αλλαγές μόλις ψηφιζόταν.
14. Πάνω στο όρος Έβερεστ, έπειτα από μέρες ταλαιπωρίας το κουνούπι σκότωσε τον ορειβάτη με ένα γερό χτύπημα.
15. Ενώ κάλπαζαν στο δάσος, καθώς έφτασαν σε ένα πεσμένο δέντρο, ο αναβάτης τρόμαξε και πέταξε κάτω το άλογο.

16. Η μπάλα χτύπησε τον παίχτη του γκολφ δυνατά και πολλοί ανησύχησαν αμέσως για το πώς θα αντιδρούσε.

17. Μέσα στη νύχτα, το όμορφο ελάφι στρίμωξε και πυροβόλησε το θηρευτή γιατί δεν τον συμπαθούσε.

18. Η συγκεκριμένη λίμνη στην Αμερική έχει πολλούς τουρίστες, αλλά και πολλούς βατράχους που τρώγονται από μύγες.

19. Μέσα στο στοιχειωμένο σπίτι επικρατούσε άκρα του τάφου σιωπή και έτσι κατάφερε ο νεαρός να τρομάξει το φάντασμα.

20. Κατά τη διάρκεια του διαγωνισμού ιππασίας, ο αναβάτης ήταν ανήσυχος και κλώτσησε ένα από τα άλογα που περίμεναν εκεί κοντά.

21. Τα παιδιά έβλεπαν και κρατούσαν σημειώσεις στο ποτάμι όταν ο διάσημος ψαράς πιάστηκε από το ψάρι.

22. Στο έρημο μονοπάτι μέσα στο δάσος, ο κυνηγός έτρεχε λαχανιασμένος να ξεφύγει από την πάπια.

23. Ο νέος αστυνόμος συνελήφθη από τον πολίτη νωρίς τα ξημερώματα, όταν τον κατέδωσαν κάποιοι γνωστοί πληροφοριοδότες.

24. Ενώ πήγαινε να σώσει την πριγκίπισσα από τον ψηλό της πύργο, ο πρωταγωνιστής πρίγκιπας σκοτώθηκε από τον δράκο.

25. Στο μάθημα της ιστορίας διαβάζαμε πώς πολλοί πειρατές λεηλατούνταν από άγρια χωριά και χαιρόμασταν που δε συμβαίνει σήμερα αυτό.

26. Όπως πρόβλεψαν οι επιστήμονες της Νάσα, η γη πέρασε τον κομήτη στις δύο και πενήντα ακριβώς το μεσημέρι της Κυριακής.

27. Η γκρι ριγέ γάτα των γειτόνων τρόμαξε όταν γραντζουνήστηκε απότομα από μια γυναίκα περαστική.

28. Η νεαρή κοπέλα ξύπνησε με έκπληξη όταν τη χάιδεψε το κουνέλι της για να της τραβήξει την προσοχή.

29. Ο διάσημος κομμωτής κουρεύτηκε από το μικρό κοριτσάκι διότι δεν είχε λεφτά να πάει σε κανονικό κομμωτήριο.

30. Νωρίς κάθε πρωί, πριν ξυπνήσουν όλα τα μέλη της οικογένειας, ο σκύλος έβαζε το λουρί του και πήγαινε βόλτα το παιδί.

31. Έπειτα από μία συναρπαστική ερμηνεία του διάσημου θεατρικού έργου στο θέατρο Μπάντμιντον, ο θεατής χειροκροτήθηκε από τους ηθοποιούς.

32. Το παιχνίδι ήταν ήδη πέραν του ημιχρόνου όταν ο κορυφαίος ποδοσφαιριστής του αγώνα σφύριξε στο διαιτητή λόγω φάουλ.

Completely Implausible/Non-reversible sentences (Long) [-1.8ex]

1. Μέσα στην κουζίνα για μεγαλύτερη καθαριότητα ο μάγειρας φορέθηκε από την κόκκινη ποδιά και ξεκίνησε το μαγείρεμα.

2. Για να είναι εγκυημένη μια καλή συγκομιδή, πρέπει πάντα οι αγρότες να φυτεύονται από τις ελιές το καλοκαίρι.

3. Αναρωτιόμασταν όλοι γιατί το τυρί έφαγε τον αρουραίο κάτω από το τραπέζι.

4. Ανήμερα των Χριστουγέννων έμαθε ο μικρός ότι ο σκύλος θάφτηκε κάπου στον κήπο από το κόκαλο.

5. Ένας αμφιλεγόμενος δημοσιογράφος γράφτηκε από ένα μικρό άρθρο κάνοντας πολλούς να το πάρουν περισσότερο στα σοβαρά.

6. Στο πρώτο του ραντεβού με την κοπέλα του, το ανέκδοτο διηγήθηκε τον αγαπημένο του κωμικό για να σπάσει ο πάγος.

7. Ο Μάρκος χάρηκε πολύ όταν ο παλιός υδραυλικός φτιάχτηκε από το σωλήνα και άρχισε να κυλάει πάλι νερό.

8. Ο αγώνας είχε προπονηθεί σκληρά για να κερδίσει τον αθλητή, πράγμα που τελικά κατάφερε.

9. Κάθε μέρα όλη μέρα το χορτάρι μασούσε την αγελάδα και συνέχιζε την ζωή του.

10. Η Άννα παρακαλούσε τη μουσική να βάλει τον ντιτζέι που της άρεσε τελευταία για να χορέψει.

11. Χθες το πρωί, αφότου ήπιε τον καφέ της, η γραμματέας γράφτηκε από την επιστολή.

12. Όλοι καμάρωναν καθώς έβλεπαν τον ζωγράφο να βάφεται από το έργο τέχνης του.

13. Το ξύλο του παλιού σπιτιού δάγκωσε τον τερμίτη γιατί του είχε σπάσει τα νεύρα.

14. Στο χαμηλόφωτο κοτέτσι, κάθε μέρα γεννούσε το αυγό μιαν κότα, αλλά σε διαφορετικές ώρες κάθε φορά.

15. Μαζεύοντας υλικά από όλη τη γύρω περιοχή, το μυρμήγκι χτίστηκε γρήγορα από τη φωλιά.

16. Όταν το παιδί έσπασε το χέρι του, πήγε στο νοσοκομείο και η ακτινογραφία έβγαλε το γιατρό.

17. Αφότου έφτασε στο απόμερο νησί, ο πειρατής θάφτηκε από τον θησαυρό προκειμένου να τον ξαναθάψει σε πιο ασφαλείς καιρούς.

18. Παλιά που ήμασταν μικροί, έπρεπε συχνά οι ενέσεις να μας κάνουν τις νοσοκόμες για να μην αρρωστήσουμε άσχημα.

19. Οι βοηθοί έφριξαν όταν το δόντι τράβηξε τον καημένο οδοντίατρο μέσα στο στόμα της Ελένης.

20. Το τηλεοπτικό ντοκιμαντέρ έδειξε πώς το δέντρο έφαγε τον κάστορα ζωντανό γιατί πεινούσε.

21. Ο Αντρέας παρακολουθούσε από περιέργεια καθώς ο επιχειρηματίας εστάλη από το φάκελο δίχως γραμματόσημο.

22. Ο καθηγητής κοίταζε από το παράθυρό του την μπουλντόζα που μαζευόταν από το χώμα γιατί έκανε θόρυβο και τον αποσπούσε.

23. Η εκπομπή έδειχνε πώς οι μοδίστρες ράβονταν από τις φούστες σε κάθε ιστορική περίοδο.

24. Το αεροπλάνο πετούσε τον πιλότο εύκολα και με στυλ και οι ταξιδιώτες χειροκροτούσαν επιδοκιμαστικά.

25. Στους Ολυμπιακούς Αγώνες του Πεκίνου, ο πρωταθλητής πετάχτηκε σε μεγάλη απόσταση από το ακόντιο και το κοινό ζητωκραύγασε.

26. Η ανατολή του ήλιου φωτογράφισε το νεαρό ρομαντικό για να το δείξει στην αγαπημένη του.

27. Ο δήμος έβγαλε ανακοίνωση πως θα κλείσει την περιοχή για να ασφαλτοστρωθεί ο φορτηγατζής από το δρόμο.

28. Πριν το Πάσχα η οικογένεια έπρεπε να καλέσει τη στέγη να επισκευάσει το χτίστη προτού φιλοξενήσουν κόσμο.

29. Κάτω από την τόση πίεση που ασκούσε, ο καθρέφτης έσπασε τη γυναίκα σε μυριάδες κομμάτια.

30. Φέτος, όπως και κάθε χρονιά, πολλοί άνθρωποι στολίστηκαν από τα Χριστουγεννιάτικα δέντρα τους ήδη από τις αρχές Νοεμβρίου.

31. Προκειμένου να εντυπωσιάσει τον κριτικό φαγητού στο ρεστοράν της, η σούπα μαγείρεψε το διάσημο σεφ της.

32. Προς μεγάλη έκπληξη όλων στο βραδινό τραπέζι, ο άνδρας δαγκώθηκε στο χέρι από τη δήθεν μικρή και αθώα τάρτα λεμονιού.

B.2 Experiment 2: P-Stranding under Sluicing in the context of Islands; Stimuli

B.2.1 Experimental Items

To save space only the first item will be presented in all 4 conditions, with all subsequent items being presented only in their Non-Sluicing, P-pied-piping version.

1. a. **Condition 1: Non-Sluicing, P-Pied-Piping**
   Στις ειδήσεις είπαν πως η αστυνομία συνέλαβε μια γυναίκα που είχε συσχετιστεί με κάποιους τρομοκράτες, αλλά δεν πληροφορηθήκαμε με ποιους τρομοκράτες συνέλαβε η αστυνομία μια γυναίκα που είχε συσχετιστεί με.

   b. **Condition 2: Non-Sluicing, P-Stranding**
   Στις ειδήσεις είπαν πως η αστυνομία συνέλαβε μια γυναίκα που είχε συσχετιστεί με κάποιους τρομοκράτες, αλλά δεν πληροφορηθήκαμε ποιους τρομοκράτες συνέλαβε η αστυνομία μια γυναίκα που είχε συσχετιστεί με.

   c. **Condition 3: Sluicing, P-Pied-Piping**
   Στις ειδήσεις είπαν πως η αστυνομία συνέλαβε μια γυναίκα που είχε συσχετιστεί με κάποιους τρομοκράτες, αλλά δεν πληροφορηθήκαμε με ποιους τρομοκράτες.

   d. **Condition 4: Sluicing, P-Stranding**
   Στις ειδήσεις είπαν πως η αστυνομία συνέλαβε μια γυναίκα που είχε συσχετιστεί με κάποιους τρομοκράτες, αλλά δεν πληροφορηθήκαμε ποιους τρομοκράτες.

2. Ο Δημήτρης σκόπευε να χρησιμοποιήσει δεδομένα ομιλητών που προέρχονταν μόνο από συγκεκριμένους τόπους της Ελλάδας, αλλά λόγω ανταγωνισμού δεν ήθελε να μαθευτεί από ποιους τόπους σκόπευε να χρησιμοποιήσει δεδομένα ομιλητών που προέρχονταν.

3. Στη μητέρα μου φαινόταν πολύ αστείο ότι η θεία τσαντιζόταν απίστευτα με κάποιους πελάτες, ωστόσο δε θυμόμασταν με ποιους πελάτες φαινόταν πολύ αστείο στη μητέρα μου ότι η θεία τσαντιζόταν.

4. Για την εργασία της η αδερφή μου ήθελε να αναλύσει ένα βίντεο που να παίρνουν συνέντευξη από κάποιον, αλλά δεν είχε αποφασίσει ακόμη από ποιον ήθελε να αναλύσει ένα βίντεο που να παίρνουν συνέντευξη.

5. Για την εκπομπή της η παρουσιάστρια έψαχνε άτομα που να έχουν ειδικευτεί σε συγκεκριμένους επιστημονικούς κλάδους, αλλά μόνο την τελευταία στιγμή μας ενημέρωσε σε ποιους κλάδους έψαχνε άτομα που να έχουν ειδικευτεί.

6. Στην καφετέρια ο Μάρκος παρατηρούσε μια κοπέλα που φαινόταν να ανυπομονεί για κάποιον, αλλά δεν ήταν ξεκάθαρο ακόμη για ποιον παρατηρούσε μια κοπέλα που φαινόταν να ανυπομονεί.

7. Στο ραδιόφωνο έψαχναν άτομα που να μπορούν να συμπαρασταθούν σε κάποιους ηλικιωμένους αρρώστους, αλλά δεν άφησαν να εννοηθεί σε ποιους έψαχναν άτομα που να μπορούν να συμπαρασταθούν .

8. Τη Δευτέρα ο θείος μας πήγε σε μια γιορτή που βασιζόταν σε κάποιον Ινδικό θεσμό, αλλά δεν μπορούσαμε καν να μαντέψουμε σε ποιον θεσμό μας πήγε σε μια γιορτή που βασιζόταν.

9. Η αστυνομία τις προάλλες έψαχνε έναν άνδρα που καυχιόταν στο γωνιακό μπαρ για έναν φόνο, αλλά δεν ακούσαμε για ποιον φόνο έψαχνε η αστυνομία έναν άνδρα που καυχιόταν.

10. Το ότι οι μπαλαρίνες επίσης εξασκούνται σε πολλούς άλλους χορούς είναι αδιαμφισβήτητο, ωστόσο δεν είμαι σίγουρος σε ποιους άλλους χορούς είναι αδιαμφισβήτητο το ότι οι μπαλαρίνες εξασκούνται.

11. Το ότι η νοσηλεύτρια διαμαρτυρόταν καθημερινά για κάποιους ήταν γνωστό, αλλά δεν μπορώ να θυμηθώ για ποιους ήταν γνωστό το ότι η νοσηλεύτρια διαμαρτυρόταν καθημερινά.

12. Πιθανολογούνταν ότι η διάσημη ηθοποιός είχε βγει με κάποιους από το χώρο της δημοσιογραφίας, αλλά η τηλεπαρουσιάστρια είχε ξεχάσει με ποιους ακριβώς πιθανολογούνταν ότι η ηθοποιός είχε βγει.

13. Το ότι η κομμώτρια αστειευόταν συχνά με κάποιον όμορφο πελάτη της ήταν κοινό μυστικό, και όλοι γνώριζαν ακριβώς με ποιον πελάτη ήταν κοινό μυστικό το ότι η κομμώτρια αστειευόταν συχνά.

14. Το ότι η σχεδιάστρια ήταν κρυφά ερωτευμένη με κάποιον ήταν φως φανάρι για τη μητέρα της, αλλά ο πατέρας της δεν είχε αντιληφθεί ακόμη με ποιον ήταν φως φανάρι ότι η σχεδιάστρια ήταν κρυφά ερωτευμένη.

15. Στο χώρο της μόδας ήταν κοινό μυστικό ότι το μοντέλο γκρίνιαζε μια ζωή για κάποιον, αλλά το ευρύ κοινό δε γνώριζε για ποιον ήταν κοινό μυστικό ότι το μοντέλο γκρίνιαζε.

16. Ήταν εμφανές ότι η αρχαιολόγος είχε κατενθουσιαστεί με έναν συγκεκριμένο αρχαιολογικό χώρο, ωστόσο οι γονείς της δεν καταλάβαιναν με ποιον χώρο ήταν εμφανές ότι η αρχαιολόγος είχε κατενθουσιαστεί.

17. Για τη νέα μελέτη, οι ερευνητές έψαχναν είδη υπό εξαφάνιση που κατάγονταν από συγκεκριμένους τόπους, αλλά δεν ήθελαν αν μαθευτεί παραέξω από ποιους τόπους έψαχναν είδη που κατάγονταν.

18. Ήταν πολύ πιθανόν ότι η πασίγνωστη εταιρεία ευθυνόταν για πολλούς οικολογικούς συναργερμούς των τελευταίων ετών, αλλά το ευρύ κοινό δεν ήξερε για ποιους συναγερμούς ήταν πολύ πιθανόν ότι η εταιρεία ευθυνόταν.

19. Είχε μεταδοθεί στο δελτίο ότι η εγκληματίας στο δικαστήριο κατηγορούνταν για πολλούς φόνους, αλλά προσωπικά δεν άκουσα για ποιους φόνους είχε μεταδοθεί ότι η εγκληματίας κατηγορούνταν.

20. Το ότι η δικηγόρος εργαζόταν για κάποιους στη μαφία ήταν βεβαίως γνωστό, αλλά κανείς δεν τολμούσε να ξεστομίσει για ποιους ακριβώς ήταν γνωστό ότι η δικηγόρος εργαζόταν.

21. Είναι πιθανόν ότι η ανηψιά της Γωγώς αρπάχτηκε με κάποιον από το σχολείο της, αλλά η Γωγώ δεν ήξερε με ποιον είναι πιθανόν ότι η ανηψιά της αρπάχτηκε.

22. Για την Αντωνία ήταν ξεκάθαρο ότι η ζηλιάρα της παρέας ήταν θυμωμένη από χθες με κάποιους, αλλά οι υπόλοιπες κοπέλες δεν είχαν ιδέα με ποιους ήταν ξεκάθαρο ότι ήταν θυμωμένη.

23. Στο φυλλάδιο οδηγιών έγραφε πως το φάρμακο ενδείκνυται για άτομα που πάσχουν από κάποιον μύκητα, αλλά δεν ήταν ξεκάθαρο από ποιον μύκητα ενδείκνυται το φάρμακο για άτομα που πάσχουν.

24. Στο δελτίο ειδήσεων μεταδόθηκε ότι πολλές χώρες διαμαρτύρονταν για κάποιον ευρωπαϊκό κανονισμό, αλλά δεν είμαι σίγουρος για ποιον κανονισμό μεταδόθηκε ότι πολλές χώρες διαμαρτύρονταν.

25. Ήταν ενοχλητικό για τους εργαζόμενους ότι ο διευθυντής της εταιρείας αναφερόταν στις ομιλίες του μόνο σε ορισμένους, και όλοι ξέραμε ακριβώς σε ποιους ήταν ιδιαίτερα ενοχλητικό ότι ο διευθυντής αναφερόταν μόνο.

26. Στην ταινία είχε γίνει αντιληπτό από τους εχθρούς ότι ο μυστικός πράκτορας μιλούσε σε κάποιον, αλλά δε γνώριζαν οι θεατές σε ποιον είχε γίνει αντιληπτό ότι ο πράκτορας μιλούσε.

27. Οι επιστήμονες έψαχναν ένα έντομο που για την αναπαραγωγή του εξαρτιόταν από κάποιον μικροοργανισμό, αλλά δεν ήταν ξεκάθαρο από ποιον μικροοργανισμό έψαχναν ένα έντομο που εξαρτιόταν.

28. Η φιλενάδα μας αγανακτισμένη έψαχνε την κοπέλα που δάνειζε τις σημειώσεις της σε κάποιον, αλλά οι υπόλοιποι δεν είχαμε καταλάβει σε ποιον έψαχνε την κοπέλα που δάνειζε τις σημειώσεις της.

29. Ήταν αρκετά πιθανόν ότι η επιτυχία της μεγάλης τραγουδίστριας εξαρτιόταν στενά από κάποιον στο χώρο του θεάματος, αλλά δεν ήταν εμφανές από ποιον ήταν πιθανόν ότι η επιτυχία της εξαρτιόταν στενά.

30. Η πυροσβεστική έψαχνε μια γάτα που είχε κρυφτεί ψηλά σ'ένα δέντρο από κάποιους, αλλά κανείς δεν ήταν σίγουρος από ποιους έψαχνε η πυροσβεστική μια γάτα που είχε κρυφτεί.

31. Το βιβλίο μιλούσε για μία όμορφη γυναίκα που αναγκάστηκε να φυλάγεται από κάποιους στη βασιλική αυλή, αλλά δεν είχε αποκαλυφθεί ακόμη από ποιους μιλούσε το βιβλίο για μία γυναίκα που αναγκάστηκε να φυλάγεται.

32. Ο νταής του προαυλίου έψαχνε την κοπέλα που έπεσε χθες το μεσημέρι πάνω σε κάποιον φίλο του, αλλά κανείς δεν είχε καταλάβει πάνω σε ποιον φίλο του έψαχνε την κοπέλα που έπεσε.

B.3 Experiment 3: P-Stranding under Contrast Sluicing; Stimuli

B.3.1 Experimental Items

To save space only the first item will be presented in all 4 conditions, with all subsequent items being presented only in their Non-Sluicing, P-pied-piping version.

1. a. **Condition 1: Non-Sluicing, P-Pied-Piping**
   Η πωλήτρια μιλούσε άσχημα για το Γιώργο, και παραξενευόμουν για ποιον άλλον μιλούσε άσχημα.

   b. **Condition 2: Non-Sluicing, P-Stranding**
   Η πωλήτρια μιλούσε άσχημα για το Γιώργο, και παραξενευόμουν ποιον άλλον μιλούσε άσχημα για.

   c. **Condition 3: Sluicing, P-Pied-Piping**
   Η πωλήτρια μιλούσε άσχημα για το Γιώργο, και παραξενευόμουν για ποιον άλλον.

   d. **Condition 4: Sluicing, P-Stranding**
   Η πωλήτρια μιλούσε άσχημα για το Γιώργο, και παραξενευόμουν ποιον άλλον.

2. Η φοιτήτρια συγκατοικεί από πέρυσι με το Γεράσιμο, αλλά δεν θυμάμαι με ποιον άλλον συγκατοικεί.

3. Η ντροπαλή της παρέας ξαφνιάστηκε στην καφετέρια με τον Έκτορα, και αναρωτιέμαι με ποιον άλλον ξαφνιάστηκε.

4. Η νονά προσευχόταν κάθε βράδυ για το Μίλτο, και αναρωτιόμουν για ποιον άλλον προσευχόταν.
5. Η γριά εξαρτιόταν σε πολλά από τον Τάκη, αλλά αναρωτιόμουν από ποιον άλλον εξαρτιόταν.
6. Η μητέρα του καυχιόταν συνεχώς για τον Αριστείδη, και θα ήθελα να ξέρω για ποιον άλλον καυχιόταν.
7. Οι δύο κοπέλες τσακώθηκαν άσχημα για τον Αντώνη, αλλά δεν είχα συναίσθηση για ποιον άλλον τσακώθηκαν.
8. Η παιδαγωγός έπαιζε στον κήπο με το Δημήτρη, αλλά δεν είδα με ποιον άλλον έπαιζε.
9. Η απουσιολόγος είχε ενημερωθεί προφανώς για το Ματθαίο, αλλά δεν ξέραμε για ποιον άλλον είχε ενημερωθεί.
10. Η μαθήτρια έμοιαζε ολοφάνερα με το Λευτέρη, αλλά παραξενευόμουν με ποιον άλλον έμοιαζε.
11. Η διευθύντρια αναφέρθηκε εκτενώς στην ομιλία της στον Ανέστη, αλλά δεν άκουσα σε ποιον άλλον αναφέρθηκε.
12. Η εισαγγελέας έπρεπε να αποφανθεί για τον Πέτρο, αλλά δεν είχα συναίσθηση για ποιον άλλον έπρεπε να αποφανθεί.
13. Η γειτόνισσα αναγκάστηκε να δανειστεί λεφτά από το Θωμά, αλλά δεν πληροφορηθήκαμε από ποιον άλλον αναγκάστηκε να δανειστεί.
14. Η καλλιτέχνιδα ήθελε να δώσει ένα δώρο στον Αποστόλη, αλλά δεν είχε αποφασίσει σε ποιον άλλον ήθελε να δώσει δώρο.
15. Η κόρη μου είπε ότι ήθελε να διαφοροποιηθεί οπωσδήποτε από τον Μενέλαο, αλλά δεν μας είπε από ποιον άλλον ήθελε να διαφοροποιηθεί.
16. Η υποκρίτρια πήγαινε κρυφά κάποιες φορές στον Αγαμέμνονα, αλλά δεν πληροφορήθηκε κανείς σε ποιον άλλον πήγαινε.
17. Η αδερφή μου ήταν θυμωμένη σίγουρα με το Δημήτρη, αλλά δεν ήθελε να με ενημερώσει με ποιον άλλον ήταν θυμωμένη.
18. Η νοσοκόμα απευθυνόταν θυμωμένα στο Γιάννη, όμως δεν μπορούσα να δω σε ποιον άλλον απευθυνόταν.
19. Η ερευνήτρια απομακρύνθηκε σίγουρα από τον Αλκιβιάδη, όμως αναρωτιέμαι από ποιον άλλον απομακρύνθηκε.

20. Η ξαδερφούλα παραπονιόταν συνεχώς για το Μάνο, αλλά δεν άκουσα για ποιον άλλον παραπονιόταν.

21. Η ξαδέρφη του χρειάστηκε να απολογηθεί επανειλημμένα για το Γρηγοράκη, αλλά δεν μπορώ να μαντέψω για ποιον άλλον χρειάστηκε να απολογηθεί.

22. Η επιμελήτρια κακοφάνηκε εμφανέστατα στο Φώτη, αλλά δεν είδα σε ποιον άλλον κακοφάνηκε.

23. Η δασκάλα στεναχωρήθηκε χθες με το Χαράλαμπο, αλλά δεν κατάλαβα με ποιον άλλον στεναχωρήθηκε.

24. Η ιδιοκτήτρια του καταστήματος γελούσε συνεχώς με τον Κωνσταντίνο, αλλά δεν ήθελε να μαθευτεί με ποιον άλλον.

25. Η καθηγήτρια αγχώθηκε στις εξετάσεις για το Μανόλη, αλλά δεν κατάλαβα για ποιον άλλον αγχώθηκε.

26. Η προδότρια είχε συσχετιστεί υπόπτως με τον Οδυσσέα, αλλά δε μαθεύτηκε με ποιον άλλον είχε συσχετιστεί.

27. Η τουρίστρια φυλαγόταν καχύποπτα από τον Ιάκωβο, αλλά δεν μας πληροφόρησε από ποιον άλλον φυλαγόταν.

28. Η διακοσμήτρια αστειευόταν εκείνο το μεσημέρι με το Νικήτα, αλλά δεν ήταν ξεκάθαρο με ποιον άλλον αστειευόταν.

29. Η παρουσιάστρια πήρε συνέντευξη σήμερα από το Φοίβο, αλλά δεν άκουσα από ποιον άλλον πήρε συνέντευξη.

30. Η νταντά ευθύνεται τα μεσημέρια για τον Κωστάκη, αλλά δεν ξέρω για ποιον άλλον ευθύνεται.

31. Η αρραβωνιαστικιά του ανυπομονούσε προφανώς για τον Σωκράτη, αλλά δεν μπορούσα να μαντέψω για ποιον άλλον ανυπομονούσε.

32. Η γιαγιά συμπαραστέκεται πάντα στον Ιάσονα, όμως αναρωτιέμαι σε ποιον άλλον συμπαραστέκεται έτσι.

33. Η θεία έπρεπε να εξηγήσει την ιστορία πάλι στον Κίμωνα, αλλά δεν μπορώ να σκεφτώ σε ποιον άλλον έπρεπε να την εξηγήσει.

34. Η σερβιτόρα άρεσε φανερά στον Κώστα, αλλά δεν παρατήρησα σε ποιον άλλον άρεσε.

35. Η λογίστρια εκνευρίστηκε αμέσως με τον Παναγιώτη, και θα ήθελα να μάθω με ποιον άλλον εκνευρίστηκε.

36. Η αστυνόμος επιτέθηκε με θάρρος στον Μαυροτρίχη, αλλά δεν έμαθα σε ποιον άλλον επιτέθηκε.

37. Η κοπέλα του αναγκάστκε να εξηγηθεί εντέλει στο Μάρκο, αλλά δεν κατάλαβα σε ποιον άλλον αναγκάστηκε να εξηγηθεί.

38. Η φεμινίστρια τρελαινόταν κρυφά για τον Θανάση, αλλά δεν μπορούσα να φανταστώ για ποιον άλλον τρελαινόταν.

39. Η αδερφούλα του κρυβόταν όλο το πρωί από τον Σωτήρη, και αναρωτιέμαι από ποιον άλλον κρυβόταν.

40. Η προπονήτρια διαφώνησε έντονα με τον Κυριάκο, όμως δεν μπόρεσα να ακούσω με ποιον άλλον διαφώνησε.

B.4 Further Information on Dative Alternation in Modern Greek

Although the double object and NP + PP frames appear to alternate freely in certain environments (153), this is not the case for all environments, as is also the case in English. Substantial literature exists on the identity or not of these two structures in English (see Kayne (1984), Green (1974), for semantic differences; Beck and Johnson (2004) for the restitutive vs. repetitive reading *again* gives to each frame; to name but a few). Following this literature, examples (154) and (155)) below provide further anecdotal data on this alternation pattern in Greek, following Kayne (1984), with additional semantic differences following Green (1974) provided directly beneath.

(153) a. Dino to gramma sto Jiorgo.
Give.1SG *Det*.N.ACC *letter*.N.ACC *to.Det*.M.ACC *Jiorgos*.M.ACC
'I am giving the letter to George.'

b. Dino tou Jiorgou to gramma.
Give.1SG *Det*.M.GEN *Jiorgos*.M.GEN *Det*.N.ACC *letter*.N.ACC
'I am giving George the letter.'

(154) a. Paradido tou tachidromou ton fakelo.
Deliver *Det*.M.GEN *postman*.M.GEN *Det*.M.ACC *folder*.M.ACC

b. ?? I paradosi tou fakelou tou tachidromou.
Det.F.NOM *delivery*.F.NOM *Det*.M.GEN *folder*.M.GEN *Det*.M.GEN *postman*.M.GEN

c. * I paradosi tou tachidromou tou fakelou
Det.F.NOM *delivery*.F.NOM *Det*.M.GEN *postman*.M.GEN *Det*.M.GEN *folder*.M.GEN

d. * Tou tachidromou i paradosi tou fakelou.
Det.M.GEN *postman*.M.GEN *Det*.F.NOM *delivery*.F.NOM *Det*.M.GEN *folder*.M.GEN
'# The giving of the postman of the parcel' or '# The postman's giving of the parcel'.

(155) a. Paradido ton fakelo ston tachidromo.
Deliver.1SG *Det*.M.ACC *folder*.M.ACC *to.Det*.M.ACC *postman*.M.ACC

b. I paradosi tou fakelou ston tachidromo.
Det.F.NOM *delivery*.F.NOM *Det*.M.GEN *folder*.M.GEN *to.Det*.M.ACC *postman*.M.ACC

c. I paradosi ston tachidromo tou fakelou.
Det.F.NOM *delivery*.F.NOM *to.Det*.M.ACC *postman*.M.ACC *Det*.M.GEN *folder*.M.GEN

d. Tou fakelou i paradosi ston tachidromo.
Det.M.GEN *folder*.M.GEN *Det*.F.NOM *delivery*.F.NOM *to.Det*.M.ACC *postman*.M.ACC
'The giving of the parcel to the postman.'

Further semantic differences

Following Green (1974), the double object is considered to always have a semantic component of possession, namely HAVE, whereas the NP + PP frame does not. This results in two main differences: the indirect object cannot be a location in the double object frame; the semantic breakdown of each frame differs. Of interest is the first point here.

(156) a. Stelno to gramma stin Athina.
Send.1SG *Det*.N.ACC *letter*.N.ACC *to.Det*.F.ACC *Athens*.F.ACC
'Send the letter to Athens.'

b. # Stelno tis Athinas to gramma.
Send.1SG *Det*.F.GEN *Athens*.F.GEN *Det*.N.ACC *letter*.N.ACC

c. # Stelno to gramma tis Athinas.
Send.1SG *Det*.N.ACC *letter*.N.ACC *Det*.F.GEN *Athens*.F.GEN
'# Send Athens the letter.'

(157) a. Klotsao tin bala sto terma.
Kick.1SG *Det*.F.ACC *ball*.F.ACC *to.Det*.N.ACC *goal*.N.ACC
'Kick the ball into the net'

b. # Klotsao tou termatos tin bala.
Kick.1SG *Det*.N.GEN *goal*.N.GEN *Det*.F.ACC *ball*.F.ACC

c. # Klotsao tin bala tou termatos.
Kick.1SG *Det*.F.ACC *ball*.F.ACC *Det*.N.GEN *goal*.N.GEN
'# Kick the net the ball.'

(158) a. O Markos eftiakse makaronia jia ti Maria.
Det.M.NOM *Markos*.M.NOM *made*.3SG *pasta*.N.ACC.PL *for Det*.F.ACC *Maria*.F.ACC
'Marcus made pasta for Mary.'

b. O Markos eftiakse tis Marias makaronia.
Det.M.NOM *Markos*.M.NOM *made*.3SG *Det*.F.GEN *Maria*.F.GEN *pasta*.N.ACC.PL
'Marcus made Mary pasta.'

(159) a. (158a) = [Marcus's making pasta] CAUSE [BECOME [EXIST (pasta)]] for the benefit of Maria

b. (158b) = [Marcus's making pasta] CAUSE [Maria HAVE pasta]

B.5 Experiment 4: P-Stranding under Regular Sluicing (German)

B.5.1 Items

B.5.1.1 Experimental Items

To save space only the first item will be presented in all 8 conditions, with all subsequent items being presented only in their Case-matching, Non-Sluicing, P-pied-piping version.

1. a. **Condition 1: Case-Matching, Non-Sluicing, P-Pied-Piping**
Diese Parole hat offenbar bei einem bestimmten Wählerkreis besonderen Anklang gefunden und so langsam zeichnet sich auch ab, bei welchem sie besonderen Anklang gefundent hat.

   b. **Condition 2: Case-Matching, Non-Sluicing, P-Stranding**
Diese Parole hat offenbar bei einem bestimmten Wählerkreis besonderen Anklang gefunden und so langsam zeichnet sich auch ab, welchem sie bei besonderen Anklang gefundent hat.

   c. **Condition 3: Case-Matching, Sluicing, P-Pied-Piping**
Diese Parole hat offenbar bei einem bestimmten Wählerkreis besonderen Anklang gefunden und so langsam zeichnet sich auch ab bei welchem.

   d. **Condition 4: Case-Matching, Sluicing, P-Pied-Piping**
Diese Parole hat offenbar bei einem bestimmten Wählerkreis besonderen Anklang gefunden und so langsam zeichnet sich auch ab welchem.

e. **Condition 5: Case-Mismatching, Non-Sluicing, P-Pied-Piping**
Diese Parole hat offenbar bei einem bestimmten Wählerkreis besonderen Anklang gefunden und so langsam zeichnet sich auch ab, bei welcher sie besonderen Anklang gefundent hat.

f. **Condition 6: Case-Mismatching, Non-Sluicing, P-Stranding**
Diese Parole hat offenbar bei einem bestimmten Wählerkreis besonderen Anklang gefunden und so langsam zeichnet sich auch ab, welcher sie bei besonderen Anklang gefundent hat.

g. **Condition 7: Case-Mismatching, Sluicing, P-Pied-Piping**
Diese Parole hat offenbar bei einem bestimmten Wählerkreis besonderen Anklang gefunden und so langsam zeichnet sich auch ab bei welcher.

h. **Condition 8: Case-Mismatching, Sluicing, P-Stranding**
Diese Parole hat offenbar bei einem bestimmten Wählerkreis besonderen Anklang gefunden und so langsam zeichnet sich auch ab welcher.

2. Das neue Cabrio kommt bei einem Käuferkreis besonders gut an, aber es überrascht die Marktforscher, bei welchem es besonders gut ankommt.

3. Die Schuld für die großen Verluste der Bank liegt offenbar bei einem einzigen Wertpapierhändler, aber die Bank will nicht bekannt geben, bei welchem die Schuld liegt.

4. Helga hat nach der Trennung bei einem ihrer Freunde Trost gesucht und man kann sich leicht zusammenreimen, bei welchem sie Trost gesucht hat.

5. Miriam hat offenbar ihre Sitznachbarin Dorothea bei einem Lehrer verpfiffen, aber ich habe noch nicht rausgekriegt, bei welchem Lehrer sie sie verpfiffen hat.

6. Die Sekretärin von Steve Jobs hat schockierenderweise bei einem fremden Geheimdienst gearbeitet und trotzdem sollte die Presse jetzt nicht herausposaunen, bei welchem Geheimdienst sie gearbeitet hat.

7. Die Krankenschwester hat sich bei einem Besucher mit einer gefährlichen Krankheit angesteckt, aber die ärzte konnten noch nicht klären, bei welchem Besucher sie sich angesteckt hat.

8. Die Firma hat sich für die Panne nur bei einem einzigen Käufer entschuldigt und du kannst dir sicher schon denken, bei welchem Käufer sie sich für die Panne entschuldigt hat.

9. Die Redakteurin meint, dass bei einem wichtigen Politiker eine Hausdurchsuchung stattgefunden hat, aber sie liess nicht einmal anklingen, bei wem die Hausdurchsuchung stattgefunden hat.

10. Am Tag der Entführung ist das Mädchen nach der Schule offenbar bei jemandem ins Auto gestiegen, aber die Polizei rätselt noch, bei wem sie ins Auto gestiegen ist.

11. Die Verdächtige will nach dem Abendessen noch bei einem Bekannten gewesen sein, aber sie kann sich angeblich nicht daran erinnern, bei wem sie nach dem Abendessen noch gewesen ist.

12. Nach dem Brand im Haus würde Bettina sich gerne bei jemandem für die Rettung ihrer Katze bedanken, aber sie hat keine Ahnung, bei wem sie sich bedanken könnte.

13. Weil sie im Restaurant schlecht bedient worden war, wollte meine alte Großtante sich natürlich bei jemandem beschweren und es ist witzig, bei wem sie sich beschweren wollte.

14. Anke meldet sich nur, weil sie sich bei jemandem einschleimen will, und wenn du dich rüberbeugst, flüstere ich dir ins Ohr, bei wem sie sich einschleimen will.

15. Dem neuesten Gerücht zufolge soll Hannelore bei einem Kollegen übernachtet haben und warte bis du hörst, bei wem sie übernachtet haben soll.

16. Am Abend vor der Tat hat Sigrid bei jemandem von ihrer Arbeit zu Abend gegessen und die Polizei versucht jetzt zu ermitteln, bei wem sie zu Abend gegessen hat.

17. Petra hat in aller Stille mit einem Mitarbeiter des russischen Konsulats geredet und niemand sollte erfahren, mit welchem sie geredet hat.

18. Weil sie zu langsam bedient worden ist, hat meine Oma sich mit einem Kellner angelegt, aber ich konnte nicht aus ihr herausbringen, mit welchem sie sich angelegt hat.

19. Aufgrund der vielen Arbeit hat die Chefin selbst mit einem Kunden telefoniert, aber ihre Sekretärin hat nicht darauf geachtet, mit welchem die Chefin telefoniert hat.

20. Die Datenschutzbeauftragte war mit einem Wissenschaftler ganz und gar unzufrieden und dennoch wollte sie nicht zu Protokoll geben, mit welchem sie ganz und gar unzufrieden war.

21. Alle meinen, dass das Kind auf dem Foto große ähnlichkeit mit einem seiner Onkel hat und sie können sich trotzdem nicht einigen, mit welchem Onkel es große ähnlichkeit hat.

22. Die Moderatorin hat sich mit einem der Redner bekannt gemacht und es wird dich schockieren, mit welchem Redner sie sich bekannt gemacht hat.

23. Meine Mutter ist gestern mit einem Kollegen essen gegangen, aber ich bin sicher, dass du nicht erraten kannst, mit welchem Kollegen meine Mutter essen gegangen ist.

24. Friederike aus dem zweiten Studienjahr hat sich oft privat mit einem Professor getroffen, aber es ist noch ungeklärt, mit welchem Professor sie sich oft privat getroffen hat.

25. Ich soll mich ab sofort immer mit jemandem aus der Marketingabteilung absprechen, aber es ist unklar, mit wem ich mich absprechen soll.

26. Antonia wollte eigentlich ihre Reiseplanung noch mit jemandem durchgehen, aber sie konnte sich nicht entscheiden, mit wem sie sie nochmal durchgehen sollte.

27. Die Gastgeberin rechnet noch mit einem weiteren Gast und sie will uns auch unbedingt wissen lassen, mit wem sie noch rechnet.

28. Frau Oldenburg soll sich mit jemandem aus der Nachbarschaft in die Haare gekriegt haben und die meisten Nachbarn haben auch schon einen Verdacht, mit wem sie sich in die Haare gekriegt hat.

29. Anja versteht sich neuerdings ausgezeichnet mit jemandem aus ihrem Studienjahr, aber die meisten Kommilitonen finden es erschreckend, mit wem sie sich in letzter Zeit gut versteht.

30. Beim Fest hat das junge Mädchen mit jemandem getanzt und ihre Schwester sollte nicht herausfinden, mit wem sie getanzt hat.

31. Für die neue Reportage musste die Reporterin mit jemandem im Gefängnis Verbindung aufnehmen und ich hätte vorher ja nicht ahnen können, mit wem sie Verbindung aufnehmen musste.

32. Die Premierministerin muss mit jemandem aus Brüssel verhandeln, aber wir haben nocht nicht vereinbart, mit wem sie verhandeln soll.

33. Jasmin hat erzählt, dass sie nachmittages jetzt oft nach einem kleinen Jungen schaut, aber sie hat nicht erwähnt, nach welchem sie nachmittags jetzt oft schaut.

34. Die erfolgreichen Schatzsucher wollen als nächstes nach einem vergrabenen Goldschatz suchen und sie haben sich auch schon darüber geeinigt, nach welchem sie suchen wollen.

35. Rebecca hat im Sportunterricht wieder nach einem Jungen getreten und die Sportlehrerin ist eigentlich nicht darüber erstaunt, nach welchem sie getreten hat.

36. Während der Mittagspause hat eine Frau Seiffert nach einem Sachbearbeiter verlangt, aber ich habe mir nicht aufgeschrieben, nach welchem sie verlangt hat.

37. Die neue Zigarettenmarke soll nach einem berühmten Sportler benannt werde, aber die Marketingabteilung überlegt noch, nach welchem Sportler die Zigaretten benannt werden sollen.

38. Die Stute Abraxa hat beim Rennen nach einem der Reiter geschnappt und die Zuschauer konnten auch genau erkennen, nach welchem Reiter sie beim Rennen geschnappt hat.

39. Die Struktur des Merkblattes richtet sich nach irgendeinem internationalen Standard und die Sachbearbeiterin hört sich jetzt um, nach welchem Standard sich die Struktur richtet.

40. Lina sehnt sich nach einem ihrer Kindergärtner, aber sie ist zu schüchtern, um mir zuzuflüstern, nach welchem Kindergärtner sie sich sehnt.

41. Die Polizei fahndet mit Hilfe der Feuerwehr und des Katastrophenschutzes großräumig nach jemandem, aber die Medien dürfen nicht berichten, nach wem die Polizei fahndet.

42. Am Abend vor ihrem Verschwinden soll Vanessa überall nach jemandem gefragt haben und die Detektive sind sich inzwischen sicher, nach wem sie überall gefragt hat.

43. Katharina hat in den Ferien schon mal nach jemandem gesehen und ich kann mich gerne genauer darüber informieren, nach wem sie in den Ferien gesehen hat.

44. Pauline hat Sehnsucht nach jemandem aus ihrer Klasse, aber ihre Mutter soll keinen Wind davon bekommen, nach wem sie Sehnsucht hat.

45. Nora hat sich am Telefon nach jemandem erkundigt, aber ihre Schwester kann sich beim besten willen nicht ins Gedächtnis rufen, nach wem sie sich am Telefon erkundigt hat.

46. Jana hat sich in der halbdunklen Kneipe nach jemandem umgesehen und jetzt zanken sich ihre Freunde darüber, nach wem sie sich in der halbdunklen Kneipe umgesehen hat.

47. Sarah verzehrt sich jetzt schon seit einiger Zeit vor Sehnsuch nach jemandem, und dennoch weiss keiner so genau, nach wem sie sich vor Sehnsucht verzehrt.

48. Stella hat heute mit spitzen Steinen nach jemandem geworfen, aber ihre Mutter schockiert vor allem, nach wem sie mit den Steinen geworfen hat.

49. Der Messwert hängt offenbar von einem noch unbekannten Faktor ab und die Laborantin zermartert sich jetzt den Kopf darüber, von welchem der Messwert abhängt.

50. Die Agentur behauptet, dass das Publikum sich von einem ihrer grottigen Werbespots angesprochen fühlt, und da muss man sich doch wirklich fragen, von welchem sich das Publikum angesprochen fühlt.

51. Das Buch soll von einem berühmten König handeln, aber die Autorin und die Verlegerin streiten sich noch darüber, von welchem es handeln soll.

52. Angela hält von einem ihrer Mitarbeiter ganz besonders viel, aber ihr Verhalten verrät nicht, von welchem sie besonders viel hält.

53. Nach ersten Unersuchungen stammt der mumifizierte Leichnam wohl von einem ägyptischen Pharao ab und moderne genetische Tests können nun die Frage klären, von welchem Pharao sie abstammt.

54. Aufgrund von Interessenkonflikten werden wir Sigrid von einem ihrer Aufträge entbinden und wir haben auch schon ausgemacht, von welchem Auftrag wir sie entbinden werden.

55. Anne schwärmt mal wieder von einem Schauspieler, aber sie hat mir verboten auszuplaudern, von welchem Schauspieler sie schwärmt.

56. Die Studentin hält von einem ihrer Professoren recht wenig, aber sie würde nie an die große Glocke hängen, von welchem Professor sie wenig hält.

57. Valentina ist von jemandem, der schöne Gedichte schreibt, ganz angetan, aber sie möchte nicht preisgeben, von wem sie ganz angetan ist.

58. Die Abgeordnete hat offenbar von jemandem Schmiergelder erhalten, aber es hat jetzt wenig Sinn darüber zu spekulieren, von wem sie Schmiergelder erhalten hat.

59. Elfriede will sich von jemandem fernhalten, dessen schlechten Einfluss sie fürchtet, aber sie hat nicht deutlich gemacht, von wem sie sich fernhalten will.

60. Ich soll dich von jemandem grüßen, den ich zufällig im Urlaub getroffen habe, und du wirst sicher darüber staunen, von wem ich dich grüßen soll.

61. Die Firma wird sich von jemandem im Team trennen müssen und die Mitarbeitervertreter sollen jetzt aushandeln, von wem die Firma sich trennen wird.

62. Leonie hat gestern Nacht - und auch die Nächte davor - von jemandem geträumt und sie hat natürlich sofort ausposaunt, von wem sie gestern geträumt hat.

63. Auf der Videoaufnahme verabschiedet sich die Kassiererin von jemandem und glücklicherweise ist gut erkennbar, von wem sie sich verabschiedet.

64. Die Rechtsanwältin hat offenbar von irgendjemandem geheime Informationen zugesteckt bekommen, aber sie wollte noch nicht einmal andeuten, von wem sie die geheimen Informationen zugesteckt bekommen hat.

B.5.1.2 Fillers

1. Hume meinte, keine Philosophie werden uns jemals hinter die täglichen Erfahrungen führen oder uns Verhaltensregeln geben können, die anders sind als die, die wir durch unser Nachdenken über das tägliche Leben erhalten.

2. Mutter und Tochter waren in die Kirche eingesperrt gewesen und konnten über das, was draußen passiert war, kein Aussagen machen.

3. Zwischen uns entwickelte sich eine Korrespondenz, und wir trafen einander dann gelegentlich in Salzburg, London, Wien und an anderen Orten.

4. Es waren nicht so sehr die Bitten an sich, sondern die Art, wie er sie vorbrachten, die es so unüberwindlich schwierig machte, nicht nachzugeben.

5. Der Sohn wird mit einigen Knechten losreiten, wodurch sich die Gefahr vermindern und er den Befehl des halsstarrigen Alten wenigstens zur Hälfte befolgt.

6. Der Damm ist zehn Meter hoch und mit Eichen bewachsen, deren Wurzelwerk dem Eisengestänge im Beton entsprechen.

7. Kurz, ich mochte zu ihm nicht gehen, ich mochte nicht zu Kreuze kriechen und zugeben, dass er recht gehabt haben.

8. Ich schauten der alten Frau mit den Hausschuhen und dem überhängenden Mantel gerade in die Augen, war aber unfähig, ihr zu zeigen, dass ich sie sah.

9. Es war vorherzusehen, dass es früher oder später zu einem solchen Versuch, die herkömmlichen Vorstellungen über Literatur zu retten, kommen würden.

10. Eine Studentin, die am Seminar zum Thema Körbe, Einkaufsnetze, Plastiktüten teilgenommen hatten, sagten mir, dass er eine ziemlich traurige Figur gemacht habe.

11. Die Kanzlerin hat zwar verlauten lassen, dass bestimmte Kürzungen zurückgenommen werden sollen, aber dafür sollen eine zusätzliche anderswo vorgenommen werden.

12. Die Wissenschaftler untersuchte fünf Arten von Paradiesvögeln, deren Gefieder schon dem bloßen Auge besonders schwarz erscheint.

13. Neuerdings kommt zu den Studenten nun noch die Nordlicht-Jäger hinzu, die Tromsö aufgrund der guten Anbindung an den Rest der Welt und der hohen Wahrscheinlichkeit schätzen, die Aurora borealis mit eigenen Augen sehen zu können.

14. Alle übrigen Bauwerke, die den ganzen Platz ohne Lücke von drei Seiten umschloss, waren ganz niedrig, flach und schmucklos.

15. Der Fuchs hat dem müden Hasen lange gejagt, bis er ihn schließlich in die Enge treiben und fangen konnte.

16. Die Szene spielt sich in einem heruntergekommenen ehemaligen Vorzeigeviertel der Stadt ab, das jetzt von konkurrierende Gruppen beherrscht wird.

17. Längst nicht alles, was der Volksmund als Unfall bezeichnet, ist auch ein Unfall im Rechtssinn, für dem bei Angestellten die Unfallversicherung aufkommen muss.

18. Der Vorstandsvorsitzende und technische Direktor der Unternehmens, wird nicht müde zu betonen, dass autonome Autos alternativlos sind und dass wir Fahrzeuge nicht mehr allzu lange privat besitzen werden.

19. Im Juni 1741 begab sich der deutsche Arzt und Naturforscher Georg Wilhelm Steller auf einer abenteuerlichen Erkundungsreise von der sibirischen Küste in die Beringstraße und nach Alaska.

20. Wem Kindern keine lesefreundliche Umgebung schafft, der kann auch nicht erwarten, dass die Kinder sich für Bücher begeistern.

21. Die anonymen Spender haben als Pappkameraden gedient, um den Blick auf schwarze Kassen aus der siebziger Jahre zu verstellen.

22. In den Laborversuchen haben Honigbienen, deren Zuckerwasser von die Forscher mit einem Neonicotinoid versetzt wurde, die Duft-Lektion nach einem Tag vergessen.

23. Bevor das Blut an Spender weitergegeben werden kann, müssen erst den weißen Blutkörperchen herausgefiltert werden.

24. Man kann sich dieser Prozess vorstellen wie eine geschäftige Flughafenhalle, in der sich Reisende durch eine dichte Menschenmenge schieben müssen, die sie nicht durchlassen will.

25. Eine Kraft wirkt nur, solange etwas beschleunigt oder abgebremst wird, und einen Menschen kann diese Kraft nur spüren, wenn etwas langsamer oder schneller wird.

26. Das Bier ist inzwischen im ganzen Land beliebt, doch nur in in der Brauerei selbst lässt sich bei Brauereitouren eines Blickes hinter die Kulissen sowie in den angeschlossenen Bierkeller werfen.

27. Kollidierte die Zubettgehenszeit mit ein spannendes Buchkapitel, dann las man eben still und heimlich im Lichtkegel der Taschenlampe unter der Decke weiter, bis sich die Tür öffnete und die Mutter im Raum stand.

28. Der Honig schnitt bei den 105 Jungen und Mädchen am besten ab und half, dem Hustenreiz zu lindern.

29. Ich wollte das Frühstück hochbringen und wollte auch schauen, ob ich schon einem offenen Blumengeschäft finde und eine Rose für Hanna kriege.

30. Ich erklärte ihm, dass meine Studien beendet und meine künftigen Konzertengagements gesichert seien und dass ich seiner Hilfe nicht länger brauchte.

31. Der Autor und wahrscheinlich auch sein Publikum haben nicht bemerkt, dass Ottomar weit feiger dasteht, wenn ein Mädchen ihm dergestalt beschämt.

32. Bei sein erstmaliges Erscheinen im Kloster war er ja auch nicht als Träger eines diplomatischen Auftrages, sondern lediglich als Lehrer gekommen.

33. Nur ist die gesistige Schöpfung etwas, woran wir nicht so eigentlich teilnehmen können, wie manchem glaubt.

34. Bei dem Vortrag ließ auf einmal der stillschöne Jüngling, nachdem er sich durch ein Blick mit jemandem verständigt hatte, eine Grimasse in seinem Gesicht sehen.

35. In einem großen Kanu näherten sich vierundzwanzig mit Pfeil und Bogen und Holzschilden bewaffnete Männer sein Schiff.

36. Nachdem sie allen Formalitäten hinter sich hatten, wurde die Hochzeit am 30. Mai in festlichem Rahmen im Ballsaal der Sparrenburg gefeiert.

37. Wie ein halb versunkenes Schiff soll das Unterwasser-Restaurant an der Südspitze Norwegens aussehen und für ein kulinarische Erlebnisse sorgen, aber wir können noch nicht sagen wofür sonst noch.

38. Der Bergführer kennt das Gebiet wie seine Westentasche, zu jedem Gipfel hat er eine Geschichte parat, im Winter macht er Schneeschuhwanderungen und ich glaube, dass im Sommer Mountainbiketouren.

39. Man sollte messen, was sich messen läßt, und das, was sich nicht, meßbar machen, sagt Galileo Galilei in einem wichtigen Aufsatz zur Naturphilosophie.

40. Gegen Ende des Schlachts wurden die Katapulat dann nicht nur mit Felsbrocken sondern mit tausenden kleiner und kleinster Steinchen geladen.

41. Der Reporter hat einen Apfelbaum in seinem Garten gepflanzt und der streitbare Nachbar hat das gleiche gemacht, aber welchem Grund, das weiss ich nicht.

42. Von 34 Stimmen, die in der Versammlung, dem Aufsichtsgremium der Medienanstalt, abgegeben wurden, gingen neunzehn Friedrich.

43. Randolf hat getestet, was geschieht, wenn man bei den Laborexperimenten Spuren von Pestiziden in das Zuckerwasser die Bienen mengt.

44. Ausnahmslos warnen die Insektenforscher davor zu glauben, man müsse nur eine Stofflasse dem Verkehr ziehen, und die Insekten kehrten zurück.

45. An manchen Tagen haben die Erwachsene ihn mit frischen Möhren oder mit einer Gewürzgurke gekitzelt, aber ich weiss nicht was noch für Gemüse.

46. Gentests an anderen Galapagos-Riesenschildkröten haben gezeigt, dass ihre DNA teilweise mit der von George übereinstimmt, womöglich wurden als manche Artgenossen auf andere Inseln verschleppt und streifen dort noch umher.

47. So wie ein nervöser Cowboy den Finger am locker Abzug hat, könnten auch manche Gehirne leicht ein Handlungsmuster auslösen und Richtung Muskeln schicken.

48. Als sie hörte, was Nils alles erlebt hatte, beschloss sie, ein Buch darüber zu schreiben, was.

49. Tell wird nach dem Apfelschuss einem von der Umstehenden gefragt, wozu er sich einen zweiten Pfeil genommen hat und antwortet, dieser sei für den Vogt bestimmt gewesen.

50. Das Buch ist der von Tischkante heruntergallen, und ich kann es leider nicht sehen, so dass ich es auch nicht aufheben kann.

51. Arnes Mutter hat sich die Tatsache zu Nutze gemacht, dass unheimlich gerne Schokolade isst, um ihn dazu zu bringen, Hausaufgaben zu seine machen.

52. Die Apotheker hatte Mitleid mit der Obdachlosen und gab ihr die Medizin, ohne in Rechnung sie ihr zu stellen.

53. Es hatte in letzter Zeit so viele Streiks gegeben, dass niemand viel Aufhebens um den heutigen Streik gemach hat, abgesehen von den praktischen Unannehmlichkeiten, die den Streik durch ausgelöst wurden.

54. Gurken haben einen fremden der Katze Eigengeruch und annähernd eine Form, Farbe und Oberflächenstruktur wie ein Reptil.

55. Die Wissenschaftler spielten ihren Versuchshunden in mehreren Kombinationen Tonaufnahmen von verschiedenen Kommandos vor, die vor den Haltern in lobender, tadelnder oder neutraler Färbung eingesprochen worden waren.

56. Die Tiere wurden den Europäern von exzessiv gejagt, weil man ihr Fleisch und die Häute wollte; zudem galten sie als Konkurrenten für die Rinderherden der Kolonialisten.

57. Ursprünglich lebte in die Art einem sehr eng begrenzten Gebiet im Monteverde-Bergregenwald von Costa Rica, was sie sehr anfällig für Störungen machte.

58. Erst im Oktober wurde das Kreuz von Handwerkern abgenommen und zum Schmied nach Garmisch Partenkirchen gebracht.

59. Steht ein Hohlkörper unter Druck, ist die in der Hülle entstehende Zugspannung nur dann zu allen Seiten gleich groß, wenn das betreffende Objekt die Form einer Kugel hat.

60. Seit Jahrhunderten haben wir jeden Winkel der Erde erforscht und sind auf jeden vorgedrungen Kontinent, haben hinter jeden Baum, auf jeden Berg geschaut.

61. Wenn wir nicht so enden wollen wie die Dinosaurier, tun wir gut daran, uns ins Weltraum auszukennen, zu wissen, was da draußen für Gefahren von drohen.

62. Die enorme Verdichtung von Masse in einem Schwarzen Loch könnte einen Maximalpunkt erreichen, von dem sie sich aus explosionsartig wieder ausdehnt, ähnlich wie bei der Entstehung unseres Universums.

63. Im halbstündigen Takt setzt sich die Seilbahn langsam in Bewegung, und ein Panoramablick über die ganze Stadt sowie die dahinterliegenden Inseln eröffnet sich aus der Kabine.

64. Die ersten Schritte mit den klobigen Schneeschuhen gleichen zwar einer Mondlandung, aber bald hat sich das Gefühl für die Schneeschuhe eingestellt und dann freut man sich, wenn es geht bergwärts.

65. Generell gilt, dass ältere Menschen und Neugeborene in höherem Maß gefährdet sind, weil ihr Immunsystem schwächelt oder ist noch nicht voll ausgebildet.

66. Der Komponist hat dem neuen Tenor es zugemutet.

67. Welches Zimmer weißt du nicht, wo sich befindet?

68. Der Waffenhändler glaubt er, dass den Politiker bestochen hat.

69. Wen fragst du dich, ob Maria nicht kennenlernen sollte?

70. In der Mensa essen viele Studenten zu Mittag.

71. Nur sehr selten hört man den leisen, krächzenden Ruf eines Schwans.

72. Was ich wissen will, ist wen wer in dieser Affäre betrügt.

73. Ich habe dem Kunden sich selbst im Spiegel gezeigt.

74. Welche Zahnpasta hat der Zahnarzt welchem Patienten empfohlen?

75. Sie hofft, das Finanzamt hat den Betrüger überlistet.

76. Die Kommission hat Valentin zwar zum Vorwurf der Geldwäsche befragt, aber danach ist die Untersuchung im Sande verlaufen.

77. Antonias Oma hat am Schluss nachgegeben und hat ihrer Enkelin den teuersten Schokoladenosterhasen im Laden gekauft.

78. Michaela fand, dass es zwar viele Vorschriften zur Müllwiederverwertung gibt, aber dass zuwenige davon wirklich umgesetzt werden.

79. Dieses Bild ist angeblich von der Gelassenheit der Katze des Malers inspiriert, die stundenlang ruhig schnurrend aus dem Fenster sehen konnte.

80. Zwei Drittel des Landes leiden laut staatlichem Wetterdienst seit nunmehr drei Jahren darunter, dass Niederschläge ausbleiben und am stärksten betroffen ist der Nordwesten.

81. Christian fragte sich, wo er seine Schlüssel gelassen hat und wo sein Portmonnaie, während er sich panisch auf den Weg zur U-Bahn machte.

82. Nach einem seltsamen Zweikampf biss das Kaninchen der Nachbarn schließlich unseren Hund und unser Hund biss sich selbst in den Schwanz.

83. Zum Büfett steuerte Karina den Roggenauflauf, Torben den Apfelkuchen, und ihr Setter den Hundeknochen bei.

84. Autonome Fahrzeuge von heute fahren unter bestimmten optimalen Bedingungen selbst, brauchen aber ein Lenkrad und einen stets aufmerksamen Fahrer, der bei Schwierigkeiten schnell das Steuer übernimmt.

85. Vier Düsen, die seit dem Vorbeiflug am Saturn vor 37 Jahren inaktiv waren, ließen sich jetzt bei einem Test anstandslos wieder in Gang setzen.

86. Wie bei vielen anderen Papageien führte gerade auch unsere Hassliebe zu einem raschen Niedergang, denn die farbenprächtigen Vögel endeten zu häufig in Käfigen.

87. Doch in den alten Wäldern und Sümpfen des Landes lebten wohl ebenfalls Millionen Exemplare des Sittichs, der die nördlichste Verbreitung aller Papageienarten aufwies.

88. Sie können sich für eine kompetitive Karriere entscheiden - oder dafür, einem Beruf nachzugehen, der nicht so hohe Anforderungen an sie stellt, und diesen dann entspannt erledigen.

89. Die Schauspielerin fragte sich, wann wohl die neue Schminke ankommen würde, denn sie wollte die alte endlich in den Müll schmeissen.

90. Seit Elisabeth Gertrud von Sayn-Wittgenstein, letzte Namensträgerin einer Nebenlinie des Adelsgeschlechts, in den 1970er Jahren eine fragwürdige Heirat einging, ist der Name auf dem Titelmarkt.

91. Blitzschnell nimmt der Oktopus die Farbe eines giftigen Plattfischs an und schlängelt mit angelegten Armen flach über den Meeresgrund.

92. Neben ökosystemen, wie den empfindlichen alpinen Bergwiesen, werden immer öfter auch archäologische Denkmäler durch Steinmännchen beschädigt, die dort von wohlmeinenden Touristen hinterlassen werden.

93. David gehört zu einer Gruppe von Wissenschaftlern, die die Daten auswerten und in einer Fachzeitschrift publizieren sollen.

94. Sobald die Kollegin im Büro oder der Gegenüber von der U-Bahn zu Gähnen beginnt, folgt binnen kürzester Zeit eine unaufhaltsame Kettenreaktion der Müdigkeit.

95. Ein anderer Ansatz wurde deswegen in einer Studie untersucht, bei der man beobachtete, wie schwer es Probanden fällt, sich vom Gähnen anderer Leute nicht anstecken zu lassen.

96. Weil sich auch im menschlichen Gewebe leicht bewegliche elektrisch geladene Teilchen befinden, werden die über das elektrische Feld mit den Ladungen im Detektor gekoppelt.

97. Dass Strom und Magnetismus so eng zusammenhängen, nutzt der Mensch für das Kochen mit Induktionsherden, aber ich weiss leider nicht mehr für was noch.

98. Wie der Riss im Detail entsteht, wird deutlich, wenn man sich vorstellt, die Hülle des Würstchens wäre von unsichtbaren Fäden durchwoben, wie ein Stück Stoff.

99. Die vier Astronauten, die sich zur Zeit auf der Internationalen Raumstation aufhalten, haben mit einem Fidget Spinner in der Schwerelosigkeit experimentiert, aber ich weiss nicht welchem anderen Spielzeug.

100. Das wird im Nachhinein, da bin ich mir sicher, so bedeutend sein wie der erste Fisch, der sich aus dem Ozean gewagt hat.

101. Weit verbreitet ist die Vorstellung, Schwarze Löcher seien gigantische Staubsauger, die Planeten, Sterne, Galaxien, Licht, einfach alles in sich hineinsaugen.

102. Paradiesvögel zeigen nicht nur das originellste Balzverhalten der Vogelwelt, sondern können auch mit dem schwärzesten Schwarz angeben, jetzt zeigen wir ihnen, welcher Farbe noch.

103. Bewegt sich das Huhn vorwärts, sieht es verwischt – da seine Pupillen viel unbeweglicher sind als die des Menschen und zudem noch an der Seite des Kopfes sitzen.

104. Um die Position eines leckeren Wurms zu erkennen, fixieren die Vögel ihn erst mit einem Auge, wiederholen dies aus einem anderen Blickwinkel und verrechnen dann die Einzelbilder zu einem 3-D-Eindruck.

105. Zudem gibt es in dem hauseigenen Panoramakino kurze Filme zur Entstehung der Polarlichter und im Kaffee dann leckeren hausgemachten Kuchen.

106. Mit kleinem Budget, aber viel Liebe zum Detail und vor allem viel Wissen wurde hier eine ständige Ausstellung zur Entstehung der Fossilien eingerichtet.

107. Proben aus Flüssen, Bächen, Badeseen und aus der Kanalisation unter einer Klinik zeigen, dass multiresistente Erreger überall sind, aber es ist nicht klar, wie gefährlich dies für Mensch und Tier eigentlich ist.

108. Da das pflanzliche öl günstig und überall leicht zu kaufen ist, lobten es die kuwaitischen Zahnmediziner als wertvolle Alternative zu chemischen Mitteln.

109. Ob Hühnersuppe zu Recht als eines der wichtigsten Hausmittel gegen Erkältung gilt, wollten Wissenschaftler in Omaha klären.

110. Vielleicht schrieb er in einem Affenzahn aus Versehen irgend etwas, was er erst lange, nachdem er es geschrieben hatte, entdeckte.

111. Sie trug auffallend schwarze Kleidung zu ihren langen, dunklen Haaren, die sie glatt zur Seite gescheitelt hatte.

112. Es gäbe eine beträchtliche Spesenpauschale für Kleidung und Reisen, für Forschungsmaterialien, die mich auf den aktuellen Wissensstand bringen sollten und mancherlei mehr.

113. Vor ihm achteten sie überhaupt nicht auf ihre Worte, sie unterhielten sich seelenruhig darüber, wie sie einen abgemurkst oder umgenietet hatten.

114. Nach der Arbeit saßen ziemlich viel Leute in der Gaststätte, doch als sie die seltsame Gesellschaft bemerkten, verkrümelten sie sich.

115. Im Winter, inmitten einer eisigen, verschneiten Stadt zu schwimmen, eingehüllt in Dampfwolken, die bis über die Straße ziehen, das hat etwas Phantastisch-Mutiges an sich.

116. Ich wage nicht zu urteilen, was hier in der Wertschätzung ganz oben rangiert, was ganz unten, und was in der Mitte.

117. Was die alten Stadtteile so belebt und den Gang nicht ermüdend macht, ist dies: Es bietet sich mit jedem Schritt ein anderer Anblick, eine überraschende Perspektive.

118. Dass die Tage der geschlossenen Stadt gezählt waren, wurde nicht erst deutlich mit den Schießereien zwischen Aufständischen und monarchistischen Regimentern.

119. Aber es bleibt trotz allem etwas Künstliches, das nach Ersatz für lebendige Unterhaltung zwischen zwei Lebenden aussieht.

120. Bei Tisch redete er ein bisschen mit, aß ein bisschen mit und schaffte es, wenn er sich übergeben musste, bis zum Klo.

121. Girlanden von schrecklichen Sagen schlingt das Volk um den überwinder der zaubereikundigen Müller, die in entlegenen Mühlen in der Lausitz ihr düsteres Handwerk verfolgen.

122. Der Heizer hatte sich wortlos umgedreht und war wieder auf die Lokomotive geklettert, um Dampf zu machen.

123. Christoffel packte sein Gewehr und folgte dem großen Strom, aus dem Hof hinaus, durch ein zerschossenes Tor.

124. Einen Moment später fiel die Tür zu, der Riegel bewegte sich, der Schlüssel knirschte im Schloss.

125. Das Denken hat die Eigenheit, dass es nächst sich selbst am liebsten über das denkt, worüber es ohne Ende denken kann.

126. Der Kindesmörder, mit dem ich mich auf eine unheimliche Weise identifiziere, ist nun doch nicht, wie erwartet, zum Tod verurteilt worden.

127. Die altgewordene Lehrerin, steif im Garten vor ihrer Schule umhergehend, sprengte gegen Abend die Blumen.

128. Die Priester gehörten offenbar einer kirchlichen Hierarchie an, denn in der Stadt begegneten die Spanier einem Priester, der unter den Priestern eine Art Bischof war.

B.5.1.3 Thermometer Items from Featherston (2009)

Cardinal well-formedness examples from German, from most acceptable Group to least.

1. **Group A**

   i. In der Mensa essen viele Studenten zu Mittag.
   ii. Nur sehr selten hört man den leisen, krächzenden Ruf eines Schwans

2. **Group B**

   i. Welche Zahnpasta hat der Zahnarzt welchem Patienten empfohlen?
   ii. Sie hofft, das Finanzamt hat den Betrüger überlistet.

3. **Group C**

   i. Was ich wissen will, ist wen wer in dieser Affäre betrügt.
   ii. Ich habe dem Kunden sich selbst im Spiegel gezeigt.

4. **Group D**

   i. Der Komponist hat dem neuen Tenor es zugemutet.====
   ii. Welches Zimmer weißt du nicht wo sich befindet?

5. **Group E**

   i. Der Waffenhändler glaubt er, dass den Politiker bestochen hat.
   ii. Wen fragst du dich, ob Maria nicht kennenlernen sollte?

B.5.2 Additional Results Tables

B.5.2.1 Greek - German Interaction Comparison Tables

Table B.1: Greek & German Three-Way Interactions: Case-Matching*P-Stranding*Sluicing

Explanatory Notes: 1) Three-way Interaction between Case-matching, P-Stranding and Sluicing, broken down by contrasting the variable levels defined in 'Contrast' column, showing z-ratio and p-values as calculated by the emmeans package in R using Tukey adjustment for multiple comparisons; mean differences calculated using estimated marginal means. 2) To save space, 'Non-sluicing' conditions have been renamed 'Overt'.

Case-Matching	P-Stranding	Sluicing	Contrast	Mean Diff.	SEM	z-ratio	p-value
.	P-Stranding	Overt	Case Mismatching - Case Matching	-0.39	0.06	6.52	$< .0001$
.	P-Pied-Piping	Overt	Case Mismatching - Case Matching	-3.44	0.06	58.44	$< .0001$
.	P-Stranding	Sluicing	Case Mismatching - Case Matching	-0.92	0.06	15.14	$< .0001$
.	P-Pied-Piping	Sluicing	Case Mismatching - Case Matching	-3.59	0.06	60.81	$< .0001$
Case Matching	.	Overt	P-Pied-Piping - P-Stranding	3.79	0.10	37.60	$< .0001$
Case Mismatching	.	Overt	P-Pied-Piping - P-Stranding	0.74	0.10	7.28	$< .0001$
Case Matching	.	Sluicing	P-Pied-Piping - P-Stranding	2.00	0.12	16.41	$< .0001$
Case Mismatching	.	Sluicing	P-Pied-Piping - P-Stranding	-0.66	0.12	5.42	$< .0001$
Case Matching	P-Stranding	.	Sluicing - Overt	1.95	0.11	18.11	$< .0001$
Case Mismatching	P-Stranding	.	Sluicing - Overt	1.41	0.11	13.06	$< .0001$
Case Matching	P-Pied-Piping	.	Sluicing - Overt	0.16	0.08	2.13	$< .335$
Case Mismatching	P-Pied-Piping	.	Sluicing - Overt	0.01	0.08	0.19	$< .9999$

Table B.2: Greek & German Three-Way Interactions: Language*Case-Matching*Sluicing

Explanatory Notes: 1) Three-way Interaction between Language, Case-matching and Sluicing, broken down by contrasting the variable levels defined in 'Contrast' column, showing z-ratio and p-values as calculated by the emmeans package in R using Tukey adjustment for multiple comparisons; mean differences calculated using estimated marginal means. 2) To save space, 'Non-sluicing' conditions have been renamed 'Overt'.

Language	Case-Matching	Sluicing	Contrast	Mean Diff.	SEM	z-ratio	p-value
German	.	Overt	Case-Mismatching - Case-Matching	-1.84	0.06	33.081	$< .0001$
German	.	Sluicing	Case-Mismatching - Case-Matching	-1.69	0.06	30.24	$< .0001$
Greek	.	Overt	Case-Mismatching - Case-Matching	-1.98	0.06	32.258	$< .0001$
Greek	.	Sluicing	Case-Mismatching - Case-Matching	-2.82	0.06	44.678	$< .0001$
German	Case-Matching	.	Sluicing - Overt	0.62	0.08	8.081	$< .0001$
German	Case-Mismatching	.	Sluicing - Overt	0.78	0.08	10.097	$< .0001$
Greek	Case-Matching	.	Sluicing - Overt	1.49	0.08	19.045	$< .0001$
Greek	Case-Mismatching	.	Sluicing - Overt	0.65	0.08	8.263	$< .0001$
.	Case-Matching	Overt	Greek - German	-0.45	0.06	7.6	$< .0001$
.	Case-Mismatching	Overt	Greek - German	-0.59	0.06	9.992	$< .0001$
.	Case-Matching	Sluicing	Greek - German	0.42	0.06	7.01	$< .0001$
.	Case-Mismatching	Sluicing	Greek - German	-0.72	0.06	12.048	$< .0001$

Table B.3: Greek & German Three-Way Interactions: Language*P-Stranding*Case-Matching

Explanatory Notes: 1) Three-way Interaction between Language, P-Stranding and Case-matching, broken down by contrasting the variable levels defined in 'Contrast' column, showing z-ratio and p-values as calculated by the emmeans package in R using Tukey adjustment for multiple comparisons; mean differences calculated using estimated marginal means. 2) To save space, 'Non-sluicing' conditions have been renamed 'Overt'.

Language	P-Stranding	Case-Matching	Contrast	Mean Diff.	SEM	z-ratio	p-value
German	P-Stranding	.	Case-Mismatching - Case-Matching	-0.25	0.06	-4.561	0.0001
German	Pied-Piping	.	Case-Mismatching - Case-Matching	-3.27	0.06	-58.723	$< .0001$
Greek	P-Stranding	.	Case-Mismatching - Case-Matching	-1.05	0.06	-16.37	$< .0001$
Greek	Pied-Piping	.	Case-Mismatching - Case-Matching	-3.75	0.06	-59.972	$< .0001$
German	.	Case-Matching	Pied-Piping - P-Stranding	3.01	0.1	30.717	$< .0001$
German	.	Case-Mismatching	Pied-Piping - P-Stranding	-0.01	0.1	-0.055	$> .9999$
Greek	.	Case-Matching	Pied-Piping - P-Stranding	2.77	0.1	27.758	$< .0001$
Greek	.	Case-Mismatching	Pied-Piping - P-Stranding	0.08	0.1	0.769	> 0.9999
.	P-Stranding	Case-Matching	Greek - German	0.11	0.06	1.761	> 0.6241
.	P-Stranding	Case-Mismatching	Greek - German	-0.69	0.06	-11.631	$< .0001$
.	Pied-Piping	Case-Matching	Greek - German	-0.13	0.06	-2.27	> 0.2457
.	Pied-Piping	Case-Mismatching	Greek - German	-0.61	0.06	-10.411	$< .0001$

Table B.4: Greek & German Three-Way Interactions: Language*P-Stranding*Sluicing

Explanatory Notes: 1) Three-way Interaction between Language, P-Stranding and Sluicing, broken down by contrasting the variable levels defined in 'Contrast' column, showing z-ratio and p-values as calculated by the emmeans package in R using Tukey adjustment for multiple comparisons; mean differences calculated using estimated marginal means. 2) To save space, 'Non-sluicing' conditions have been renamed 'Overt'.

Language	P-Stranding	Sluicing	Contrast	Mean Diff.	SEM	z-ratio	p-value
German	.	Overt	Pied-Piping - P-Stranding	2.15	0.1	21.375	$< .0001$
German	.	Sluicing	Pied-Piping - P-Stranding	0.85	0.12	6.997	$< .0001$
Greek	.	Overt	Pied-Piping - P-Stranding	2.37	0.1	23.702	$< .0001$
Greek	.	Sluicing	Pied-Piping - P-Stranding	0.48	0.12	3.979	< 0.0008
German	P-Stranding	.	Sluicing - Overt	1.35	0.11	12.448	$< .0001$
German	Pied-Piping	.	Sluicing - Overt	0.05	0.08	0.644	> 0.9998
Greek	P-Stranding	.	Sluicing - Overt	2.01	0.11	18.731	$< .0001$
Greek	Pied-Piping	.	Sluicing - Overt	0.13	0.08	1.689	> 0.6824
.	P-Stranding	Overt	Greek - German	-0.63	0.06	-10.537	$< .0001$
.	Pied-Piping	Overt	Greek - German	-0.42	0.06	-7.153	$< .0001$
.	P-Stranding	Sluicing	Greek - German	0.04	0.06	0.619	> 0.9999
.	Pied-Piping	Sluicing	Greek - German	-0.33	0.06	-5.744	$< .0001$

Table B.5: Greek & German Four-Way Interaction

Explanatory Notes: 1) Four-way Interaction between Language, P-Stranding, Sluicing and Case-matching, broken down by contrasting the variable levels defined in 'Contrast' column, showing z-ratio and p-values as calculated by the emmeans package in R using Tukey adjustment for multiple comparisons; mean differences calculated using estimated marginal means. 2) To save space, 'Non-sluicing' conditions have been renamed 'Overt'; 'Case-matching' conditions 'Case-match' and 'Case-mismatching' conditions 'Case-mismatch'.

Language	P-Stranding	Sluicing	Case-Matching	Contrast	Mean Diff.	SEM	z-ratio	p-value
German	P-Stranding	Overt	.	Case-Mismatch - Case-Match	−0.5	0.08	−6.385	< .0001
German	Pied-Piping	Overt	.	Case-Mismatch - Case-Match	−3.18	0.08	−40.346	< .0001
German	P-Stranding	Sluicing	.	Case-Mismatch - Case-Match	−0.01	0.08	−0.074	> .999
German	Pied-Piping	Sluicing	.	Case-Mismatch - Case-Match	−3.37	0.08	−42.711	< .0001
Greek	P-Stranding	Overt	.	Case-Mismatch - Case-Match	−0.27	0.09	−3.045	> 0.0719
Greek	Pied-Piping	Overt	.	Case-Mismatch - Case-Match	−3.7	0.09	−42.722	< .0001
Greek	P-Stranding	Sluicing	.	Case-Mismatch - Case-Match	−1.84	0.09	−20.136	< .0001
Greek	Pied-Piping	Sluicing	.	Case-Mismatch - Case-Match	−3.81	0.09	−43.149	< .0001
German	.	Overt	Case-Match	Pied-Piping - P-Stranding	3.49	0.12	30.325	< .0001
German	.	Overt	Case-Mismatch	Pied-Piping - P-Stranding	0.82	0.12	7.085	< .0001
German	.	Sluicing	Case-Match	Pied-Piping - P-Stranding	2.53	0.13	18.885	< .0001
German	.	Sluicing	Case-Mismatch	Pied-Piping - P-Stranding	−0.83	0.13	−6.169	< .0001
Greek	.	Overt	Case-Match	Pied-Piping - P-Stranding	4.08	0.12	34.797	< .0001
Greek	.	Overt	Case-Mismatch	Pied-Piping - P-Stranding	0.65	0.12	5.557	< .0001
Greek	.	Sluicing	Case-Match	Pied-Piping - P-Stranding	1.46	0.14	10.712	< .0001
Greek	.	Sluicing	Case-Mismatch	Pied-Piping - P-Stranding	−0.5	0.14	−3.642	< 0.0087
German	P-Stranding	.	Case-Match	Sluicing - Overt	1.1	0.12	9.04	< .0001
German	P-Stranding	.	Case-Mismatch	Sluicing - Overt	1.6	0.12	13.104	< .0001
German	Pied-Piping	.	Case-Match	Sluicing - Overt	0.14	0.09	1.504	> 0.9894
German	Pied-Piping	.	Case-Mismatch	Sluicing - Overt	−0.04	0.09	−0.463	> .9999
Greek	P-Stranding	.	Case-Match	Sluicing - Overt	2.8	0.12	22.571	< .0001
Greek	P-Stranding	.	Case-Mismatch	Sluicing - Overt	1.23	0.12	9.854	< .0001
Greek	Pied-Piping	.	Case-Match	Sluicing - Overt	0.18	0.1	1.888	> 0.8573
Greek	Pied-Piping	.	Case-Mismatch	Sluicing - Overt	0.07	0.1	0.746	> .9999
.	P-Stranding	Overt	Case-Match	Greek - German	−0.74	0.08	−9.123	< .0001
.	P-Stranding	Overt	Case-Mismatch	Greek - German	−0.51	0.08	−5.921	< .0001
.	Pied-Piping	Overt	Case-Match	Greek - German	−0.15	0.08	−1.833	> 0.891
.	Pied-Piping	Overt	Case-Mismatch	Greek - German	−0.67	0.08	−8.413	< .0001
.	P-Stranding	Sluicing	Case-Match	Greek - German	0.95	0.09	10.94	< .0001
.	P-Stranding	Sluicing	Case-Mismatch	Greek - German	−0.88	0.08	−10.721	< .0001
.	Pied-Piping	Sluicing	Case-Match	Greek - German	−0.11	0.08	−1.405	> 0.996
.	Pied-Piping	Sluicing	Case-Mismatch	Greek - German	−0.55	0.08	−6.521	< .0001

Appendix C

Appendix: Chapter 4

C.1 Greek SPR Study

C.1.1 Residual RT Mean and SEM Results per Region per Condition

Table C.1: Greek SPR RT Means and SEM per Region per Condition

but.NEG. (C)

	Condition	RT Resid. Mean	SEM	Log10 Resid. Mean	SEM
1	Wh-NP; Gender Matching	-33.730	32.489	-0.068	0.062
2	Wh-NP; Gender Mismatching	-36.548	39.072	-0.070	0.065
3	Wh-PP; Gender Matching	-56.744	44.404	-0.105	0.073
4	Wh-PP; Gender Mismatching	-61.313	47.282	-0.105	0.070

was.clear (VP)

	Condition	RT Resid. Mean	SEM	Log10 Resid. Mean	SEM
1	Wh-NP; Gender Matching	-47.221	26.829	-0.101	0.061
2	Wh-NP; Gender Mismatching	-52.931	21.975	-0.109	0.041
3	Wh-PP; Gender Matching	-65.385	25.958	-0.121	0.051
4	Wh-PP; Gender Mismatching	-59.924	27.047	-0.110	0.050

for (P)

	Condition	RT Resid. Mean	SEM	Log10 Resid. Mean	SEM
1	Wh-NP; Gender Matching	N/A	N/A	N/A	N/A
2	Wh-NP; Gender Mismatching	N/A	N/A	N/A	N/A
3	Wh-PP; Gender Matching	26.505	10.617	0.051	0.022
4	Wh-PP; Gender Mismatching	45.044	10.603	0.074	0.018

which.paintings (WH)

	Condition	RT Resid. Mean	SEM	Log10 Resid. Mean	SEM
1	Wh-NP; Gender Matching	1.197	27.636	0.012	0.052
2	Wh-NP; Gender Mismatching	7.756	20.474	0.054	0.040
3	Wh-PP; Gender Matching	-46.632	23.090	-0.069	0.044
4	Wh-PP; Gender Mismatching	-59.367	26.431	-0.080	0.050

the.self.hers (Reflexive)

	Condition	RT Resid. Mean	SEM	Log10 Resid. Mean	SEM
1	Wh-NP; Gender Matching	32.988	30.859	0.073	0.057
2	Wh-NP; Gender Mismatching	38.503	21.252	0.104	0.043
3	Wh-PP; Gender Matching	4.249	35.018	-0.012	0.061
4	Wh-PP; Gender Mismatching	18.419	27.134	0.034	0.070

from.the.exhibition (Embedded PP)

	Condition	RT Resid. Mean	SEM	Log10 Resid. Mean	SEM
1	Wh-NP; Gender Matching	-104.144	18.500	-0.191	0.044
2	Wh-NP; Gender Mismatching	-18.769	23.518	-0.052	0.041
3	Wh-PP; Gender Matching	-112.307	25.860	-0.212	0.050
4	Wh-PP; Gender Mismatching	-53.543	34.497	-0.101	0.066

was.pleased (Embedded VP)

	Condition	RT Resid. Mean	SEM	Log10 Resid. Mean	SEM
1	Wh-NP; Gender Matching	39.591	30.917	0.023	0.037
2	Wh-NP; Gender Mismatching	85.795	40.161	0.097	0.050
3	Wh-PP; Gender Matching	19.322	34.680	0.017	0.041
4	Wh-PP; Gender Mismatching	71.884	38.262	0.089	0.047

the.model (Embedded NP)

	Condition	RT Resid. Mean	SEM	Log10 Resid. Mean	SEM
1	Wh-NP; Gender Matching	125.028	50.757	0.235	0.061
2	Wh-NP; Gender Mismatching	147.125	45.298	0.312	0.074
3	Wh-PP; Gender Matching	85.806	35.603	0.215	0.071
4	Wh-PP; Gender Mismatching	78.515	46.995	0.194	0.075

quietly (End)

	Condition	RT Resid. Mean	SEM	Log10 Resid. Mean	SEM
1	Wh-NP; Gender Matching	297.115	83.473	0.515	0.094
2	Wh-NP; Gender Mismatching	370.771	73.635	0.630	0.093
3	Wh-PP; Gender Matching	262.286	73.620	0.480	0.099
4	Wh-PP; Gender Mismatching	328.206	58.700	0.566	0.083

C.1.2 Raw RT Mean and SEM Results per Region per Condition

but.NEG. (C)

	Condition	RT Mean	SEM	Log10 RT Mean	SEM
1	Wh-NP; Gender Matching	973.671	50.268	12.951	0.208
2	Wh-NP; Gender Mismatching	981.053	47.640	13.057	0.139
3	Wh-PP; Gender Matching	981.837	48.371	13.286	0.293
4	Wh-PP; Gender Mismatching	972.876	44.468	13.226	0.209

was.clear (VP)

	Condition	RT Mean	SEM	Log10 RT Mean	SEM
1	Wh-NP; Gender Matching	728.022	61.207	10.365	0.449
2	Wh-NP; Gender Mismatching	747.647	60.901	10.674	0.502
3	Wh-PP; Gender Matching	724.806	65.023	10.594	0.550
4	Wh-PP; Gender Mismatching	753.192	69.180	11.003	0.684

for (P)

	Condition	RT Mean	SEM	Log10 RT Mean	SEM
1	Wh-NP; Gender Matching	N/A	N/A	N/A	N/A
2	Wh-NP; Gender Mismatching	N/A	N/A	N/A	N/A
3	Wh-PP; Gender Matching	404.396	23.297	5.931	0.054
4	Wh-PP; Gender Mismatching	423.432	25.847	5.955	0.052

which.paintings (WH)

	Condition	RT Mean	SEM	Log10 RT Mean	SEM
1	Wh-NP; Gender Matching	878.249	64.950	11.795	0.157
2	Wh-NP; Gender Mismatching	888.909	58.574	11.919	0.127
3	Wh-PP; Gender Matching	842.909	55.671	11.849	0.124
4	Wh-PP; Gender Mismatching	831.462	48.948	11.843	0.113

the.self.hers (Reflexive)

	Condition	RT Mean	SEM	Log10 RT Mean	SEM
1	Wh-NP; Gender Matching	1387.266	85.050	19.281	0.271
2	Wh-NP; Gender Mismatching	1402.534	79.100	19.423	0.287
3	Wh-PP; Gender Matching	1411.729	74.966	19.273	0.223
4	Wh-PP; Gender Mismatching	1430.151	76.602	19.333	0.246

from.the.exhibition (Embedded PP)

	Condition	RT Mean	SEM	Log10 RT Mean	SEM
1	Wh-NP; Gender Matching	1066.201	68.742	14.479	0.221
2	Wh-NP; Gender Mismatching	1170.859	68.030	14.817	0.196
3	Wh-PP; Gender Matching	1135.085	65.576	15.101	0.229
4	Wh-PP; Gender Mismatching	1205.951	66.213	15.352	0.176

was.pleased (Embedded VP)

	Condition	RT Mean	SEM	Log10 RT Mean	SEM
1	Wh-NP; Gender Matching	709.236	57.097	7.552	0.170
2	Wh-NP; Gender Mismatching	767.404	63.530	7.768	0.191
3	Wh-PP; Gender Matching	702.785	62.461	7.617	0.144
4	Wh-PP; Gender Mismatching	762.496	55.581	7.759	0.172

the.model (Embedded NP)

	Condition	RT Mean	SEM	Log10 Resid. Mean	SEM
1	Wh-NP; Gender Matching	1615.926	112.839	18.221	0.230
2	Wh-NP; Gender Mismatching	1636.932	87.063	18.251	0.198
3	Wh-PP; Gender Matching	1631.222	76.887	18.678	0.163
4	Wh-PP; Gender Mismatching	1604.062	83.785	18.423	0.183

quietly (End)

	Condition	RT Mean	SEM	Log10 RT Mean	SEM
1	Wh-NP; Gender Matching	1945.642	148.682	19.407	0.565
2	Wh-NP; Gender Mismatching	2021.531	142.359	19.582	0.574
3	Wh-PP; Gender Matching	1946.530	117.243	19.994	0.678
4	Wh-PP; Gender Mismatching	1965.198	120.110	19.627	0.714

C.1.3 Order Effect Analysis Results

C.1.3.1 LogRT Residuals Main Effects and Interactions

Table C.2: Greek SPR Linear Regression Analysis Results

	Higher Order Interactions			
	Gender-Matching*Wh-Type*Order			
Region	β	SE	t	p
for (P)	N/A	N/A	N/A	N/A
which.paintings (Wh)	0.086	0.132	0.652	0.514
the.self.hers (Reflexive)	0.062	0.187	0.333	0.739
from.the.exhibition (Embedded PP)	0.346	0.170	1.910	0.070
was.pleased (Embedded VP)	0.055	0.115	0.475	0.634
the.model (Embedded NP)	0.120	0.177	0.678	0.498
quietly (End)	0.058	0.202	0.290	0.772

	Lower Order Interactions (1)			
	Gender-Matching*WhType			
Region	β	SE	t	p
for (P)	N/A	N/A	N/A	N/A
which.paintings (Wh)	0.020	0.087	0.227	0.820
the.self.hers (Reflexive)	0.078	0.098	0.791	0.429
from.the.exhibition (Embedded PP)	0.024	0.110	0.218	0.827
was.pleased (Embedded VP)	0.025	0.073	0.340	0.734
the.model (Embedded NP)	0.060	0.098	0.612	0.541
quietly (End)	0.031	0.123	0.255	0.799

	Lower Order Interactions (2)			
	Gender-Matching*Order			
Region	beta	SE	t	p
for (P)	0.002	0.060	0.038	0.970
which.paintings (Wh)	0.061	0.066	0.922	0.357
the.self.hers (Reflexive)	0.013	0.093	0.143	0.886
from.the.exhibition (Embedded PP)	0.091	0.085	1.063	0.288
was.pleased (Embedded VP)	0.016	0.058	0.284	0.776
the.model (Embedded NP)	0.037	0.088	0.424	0.672
quietly (End)	0.067	0.101	0.667	0.505

	Lower Order Interactions (3)			
	Wh-Type*Order			
Region	beta	SE	t	p
for (P)	N/A	N/A	N/A	N/A
which.paintings (Wh)	0.084	0.066	1.272	0.203
the.self.hers (Reflexive)	0.073	0.093	0.784	0.433
from.the.exhibition (Embedded PP)	0.043	0.085	0.503	0.615
was.pleased (Embedded VP)	0.058	0.058	1.011	0.312
the.model (Embedded NP)	0.052	0.088	0.593	0.553
quietly (End)	0.015	0.100	0.145	0.885

	Main Effects (1)			
	Gender-Matching			
Region	beta	SE	t	p
for (P)	0.021	0.033	0.630	0.528
which.paintings (Wh)	0.010	0.033	0.315	0.753
the.self.hers (Reflexive)	0.002	0.058	0.039	0.969
from.the.exhibition (Embedded PP)	0.088	0.046	2.10	0.048*
was.pleased (Embedded VP)	0.043	0.034	1.261	0.207
the.model (Embedded NP)	0.003	0.054	0.050	0.960
quietly (End)	0.076	0.062	1.219	0.223

	Main Effects (2)			
	Wh-Type			
Region	beta	SE	t	p
for (P)	N/A	N/A	N/A	N/A
which.paintings (Wh)	0.082	0.048	1.687	0.092
the.self.hers (Reflexive)	0.031	0.052	0.599	0.549
from.the.exhibition (Embedded PP)	0.001	0.042	0.018	0.986
was.pleased (Embedded VP)	0.009	0.040	0.217	0.828
the.model (Embedded NP)	0.038	0.050	0.748	0.455
quietly (End)	0.015	0.058	0.260	0.795

	Main Effects (3)			
	Order			
Region	beta	SE	t	p
for (P)	0.223	0.030	7.480	$< 0.001^{***}$
which.paintings (Wh)	0.509	0.033	15.484	$< 0.001^{***}$
the.self.hers (Reflexive)	0.870	0.046	18.805	$< 0.001^{***}$
from.the.exhibition (Embedded PP)	0.785	0.042	18.547	$< 0.001^{***}$
was.pleased (Embedded VP)	0.452	0.029	15.800	$< 0.001^{***}$
the.model (Embedded NP)	0.632	0.044	14.528	$< 0.001^{***}$
quietly (End)	0.683	0.050	13.723	$< 0.001^{***}$

C.1.4 Greek SPR Stimuli

Experimental Items

1. i. Ο πατέρας της Νικολέτας εξιστόρησε πολλούς θρύλους στη θερινή γιορτή, αλλά δεν είναι ξεκάθαρο ποιον θρύλο για τον εαυτό του από τη γιορτή άκουσε ο παππούς της αργότερα και στεναχωρήθηκε.

   ii. Η μητέρα του Νικολάκη εξιστόρησε πολλούς θρύλους στη θερινή γιορτή, αλλά δεν είναι ξεκάθαρο ποιον θρύλο για τον εαυτό του από τη γιορτή άκουσε ο παππούς της αργότερα και στεναχωρήθηκε.

   iii. Ο πατέρας της Νικολέτας εξιστόρησε πολλούς θρύλους στη θερινή γιορτή, αλλά δεν είναι ξεκάθαρο για ποιον θρύλο για τον εαυτό του από τη γιορτή άκουσε ο παππούς της αργότερα και στεναχωρήθηκε.

   iv. Η μητέρα του Νικολάκη εξιστόρησε πολλούς θρύλους στη θερινή γιορτή, αλλά δεν είναι ξεκάθαρο για ποιον θρύλο για τον εαυτό του από τη γιορτή άκουσε ο παππούς της αργότερα και στεναχωρήθηκε.

2. i. Ο πατέρας της Ελένης αντιλήφθηκε τυχαία κάποιους χαρακτηρισμούς των μελών στη συνάντηση γονέων, αλλά δεν είναι ξεκάθαρο ποιον χαρακτηρισμό του εαυτού του στη συνάντηση άκουσε ο καθηγητής της κάποια στιγμή και προσβλήθηκε.

   ii. Η μητέρα του Έκτορα αντιλήφθηκε τυχαία κάποιους χαρακτηρισμούς των μελών στη συνάντηση γονέων, αλλά δεν είναι ξεκάθαρο ποιον χαρακτηρισμό του εαυτού του στη συνάντηση άκουσε ο καθηγητής του κάποια στιγμή και προσβλήθηκε.

   iii. Ο πατέρας της Ελένης αντιλήφθηκε τυχαία κάποιους χαρακτηρισμούς των μελών στη συνάντηση γονέων, αλλά δεν είναι ξεκάθαρο για ποιον χαρακτηρισμό του εαυτού του στη συνάντηση άκουσε ο καθηγητής της κάποια στιγμή και προσβλήθηκε.

   iv. Η μητέρα του Έκτορα αντιλήφθηκε τυχαία κάποιους χαρακτηρισμούς των μελών στη συνάντηση γονέων, αλλά δεν είναι ξεκάθαρο για ποιον χαρακτηρισμό του εαυτού του στη συνάντηση άκουσε ο καθηγητής του κάποια στιγμή και προσβλήθηκε.

3. i. Η μαμά του Γιώργου καταδίκασε κάποιους αυθαίρετους διορισμούς σε νομικές επιτροπές, αλλά κανείς δεν θυμήθηκε ποιον διορισμό του εαυτού της στην επιτροπή είχε ομολογήσει η αδερφή της πριν μερικά χρόνια.

   ii. Ο μπαμπάς της Γεωργίας καταδίκασε κάποιους αυθαίρετους διορισμούς σε νομικές επιτροπές, αλλά κανείς δεν θυμήθηκε ποιον διορισμό του εαυτού της στην επιτροπή είχε ομολογήσει η αδερφή της πριν μερικά χρόνια.

   iii. Η μαμά του Γιώργου καταδίκασε κάποιους αυθαίρετους διορισμούς σε νομικές επιτροπές, αλλά κανείς δεν θυμήθηκε για ποιον διορισμό του εαυτού της στην επιτροπή είχε ομολογήσει η αδερφή της πριν μερικά χρόνια.

iv. Ο μπαμπάς της Γεωργίας καταδίκασε κάποιους αυθαίρετους διορισμούς σε νομικές επιτροπές, αλλά κανείς δεν θυμήθηκε για ποιον διορισμό του εαυτού της στην επιτροπή είχε ομολογήσει η αδερφή της πριν μερικά χρόνια.

4. i. Ο μπαμπάς της Δανάης εξύμνησε ορισμένους προγόνους της οικογένειας στην οικογενειακή συγκέντρωση, αλλά δεν ήταν ξεκάθαρο ποιον πρόγονο του εαυτού του στη συγκέντρωση είχε θρηνήσει ο φίλος της σιωπηρά.

   ii. Η μαμά του Σωτήρη εξύμνησε ορισμένους προγόνους της οικογένειας στην οικογενειακή συγκέντρωση, αλλά δεν ήταν ξεκάθαρο ποιον πρόγονο του εαυτού του στη συγκέντρωση είχε θρηνήσει ο φίλος της σιωπηρά.

   iii. Ο μπαμπάς της Δανάης εξύμνησε ορισμένους προγόνους της οικογένειας στην οικογενειακή συγκέντρωση, αλλά δεν ήταν ξεκάθαρο για ποιον πρόγονο του εαυτού του στη συγκέντρωση είχε θρηνήσει ο φίλος της σιωπηρά.

   iv. Η μαμά του Σωτήρη εξύμνησε ορισμένους προγόνους της οικογένειας στην οικογενειακή συγκέντρωση, αλλά δεν ήταν ξεκάθαρο για ποιον πρόγονο του εαυτού του στη συγκέντρωση είχε θρηνήσει ο φίλος της σιωπηρά.

5. i. Ο παππούς της Στέλλας αναγνώρισε ορισμένους ευεργέτες του φιλανθρωπικού οργανισμού στην εφημερίδα, αλλά κανείς δεν κατάλαβε πόσους ευεργέτες του εαυτού της στην εκκλησία γνώριζε η διευθύντρια του οργανισμού προσωπικά.

   ii. Η γιαγιά του Χρήστου αναγνώρισε ορισμένους ευεργέτες του φιλανθρωπικού οργανισμού στην εφημερίδα, αλλά κανείς δεν κατάλαβε πόσους ευεργέτες του εαυτού της στην εκκλησία γνώριζε η διευθύντρια του οργανισμού προσωπικά.

   iii. Ο παππούς της Στέλλας αναγνώρισε ορισμένους ευεργέτες του φιλανθρωπικού οργανισμού στην εφημερίδα, αλλά κανείς δεν κατάλαβε για πόσους ευεργέτες του εαυτού της στην εκκλησία γνώριζε η διευθύντρια του οργανισμού προσωπικά.

   iv. Η γιαγιά του Χρήστου αναγνώρισε ορισμένους ευεργέτες του φιλανθρωπικού οργανισμού στην εφημερίδα, αλλά κανείς δε κατάλαβε για πόσους ευεργέτες του εαυτού της στην εκκλησία γνώριζε η διευθύντρια του οργανισμού προσωπικά.

6. i. Ο παππούς της Ματίνας είχε αναθεματίσει κάποιους σωσίες καλλιτεχνών στο φεστιβάλ τέχνης, αλλά δεν ξέραμε ποιον σωσία του εαυτού του στο φεστιβάλ έβριζε με τόσο πάθος ο καλλιτέχνης αδερφός του.

   ii. Η γιαγιά του Μιχάλη είχε αναθεματίσει κάποιους σωσίες καλλιτεχνών στο φεστιβάλ τέχνης, αλλά δεν ξέραμε ποιον σωσία του εαυτού του στο φεστιβάλ έβριζε με τόσο πάθος ο καλλιτέχνης αδερφός του.

iii. Ο παππούς της Ματίνας είχε αναθεματίσει κάποιους σωσίες καλλιτεχνών στο φεστιβάλ τέχνης, αλλά δεν ξέραμε για ποιον σωσία του εαυτού του στο φεστιβάλ έβριζε με τόσο πάθος ο καλλιτέχνης αδερφός του.

iv. Η γιαγιά του Μιχάλη είχε αναθεματίσει κάποιους σωσίες καλλιτεχνών στο φεστιβάλ τέχνης, αλλά δεν ξέραμε για ποιον σωσία του εαυτού του στο φεστιβάλ έβριζε με τόσο πάθος ο καλλιτέχνης αδερφός του.

7. i. Ο νονός της Κάλλιας τοιχοκόλλησε κάποιους νέους διορισμούς στο υπουργικό γραφείο, αλλά δεν είχε αντιληφθεί κανείς ποιο διορισμό του εαυτού του σε ανάλογο γραφείο είχε γιορτάσει ο αδερφός της νωρίτερα μέσα στην εβδομάδα.

ii. Η νονά του Πέτρου τοιχοκόλλησε κάποιους νέους διορισμούς στο υπουργικό γραφείο, αλλά δεν είχε αντιληφθεί κανείς ποιο διορισμό του εαυτού του σε ανάλογο γραφείο είχε γιορτάσει ο αδερφός της νωρίτερα μέσα στην εβδομάδα.

iii. Ο νονός της Κάλλιας τοιχοκόλλησε κάποιους νέους διορισμούς στο υπουργικό γραφείο, αλλά δεν είχε αντιληφθεί κανείς για ποιο διορισμό του εαυτού του σε ανάλογο γραφείο είχε γιορτάσει ο αδερφός της νωρίτερα μέσα στην εβδομάδα.

iv. Η νονά του Πέτρου τοιχοκόλλησε κάποιους νέους διορισμούς στο υπουργικό γραφείο, αλλά δεν είχε αντιληφθεί κανείς για ποιο διορισμό του εαυτού του σε ανάλογο γραφείο είχε γιορτάσει ο αδερφός της νωρίτερα μέσα στην εβδομάδα.

8. i. Η νονά του Άγγελου ανακοίνωσε πολλούς επαίνους στην τελική γιορτή της χρονιάς, αλλά κανείς δεν κατάλαβε ποιον έπαινο του εαυτού της στη γιορτή χάρηκε η κόρη της περισσότερο από όλους.

ii. Ο νονός της Αγγέλας ανακοίνωσε πολλούς επαίνους στην τελική γιορτή της χρονιάς, αλλά κανείς δεν κατάλαβε ποιον έπαινο του εαυτού της στη γιορτή χάρηκε η κόρη της περισσότερο από όλους.

iii. Η νονά του Άγγελου ανακοίνωσε πολλούς επαίνους στην τελική γιορτή της χρονιάς, αλλά κανείς δεν κατάλαβε για ποιον έπαινο του εαυτού της στη γιορτή χάρηκε η κόρη της περισσότερο από όλους.

iv. Ο νονός της Αγγέλας ανακοίνωσε πολλούς επαίνους στην τελική γιορτή της χρονιάς, αλλά κανείς δεν κατάλαβε για ποιον έπαινο του εαυτού της στη γιορτή χάρηκε η κόρη της περισσότερο από όλους.

9. i. Η αδερφή του Κωνσταντίνου έφερε κάποιους πίνακες στο σχολικό καρναβάλι, αλλά κανείς δεν μπορούσε να θυμηθεί πόσους πίνακες του εαυτού της από το καρναβάλι ζήλεψε η κοπέλα του αργότερα.

ii. Ο αδερφός της Κωνσταντίνας έφερε κάποιους πίνακες στο σχολικό καρναβάλι, αλλά κανείς δεν μπορούσε να θυμηθεί πόσους πίνακες του εαυτού της από το καρναβάλι ζήλεψε η κοπέλα του αργότερα.

iii. Η αδερφή του Κωνσταντίνου έφερε κάποιους πίνακες στο σχολικό καρναβάλι, αλλά κανείς δεν μπορούσε να θυμηθεί πόσους πίνακες του εαυτού της από το καρναβάλι ζήλεψε η κοπέλα του αργότερα.

iv. Ο αδερφός της Κωνσταντίνας έφερε κάποιους πίνακες στο σχολικό καρναβάλι, αλλά κανείς δεν μπορούσε να θυμηθεί πόσους πίνακες του εαυτού της από το καρναβάλι ζήλεψε η κοπέλα του αργότερα.

10. i. Η αδερφή του Αντώνη κατήγγειλε πολλούς χρηματισμούς μελών της Ολυμπιακής επιτροπής στην κυβέρνηση, αλλά δεν ανέφερε πόσους χρηματισμούς του εαυτού της από την επιτροπή είχε ομολογήσει η φίλη της μόλις την προηγούμενη ημέρα.

ii. Ο αδερφός της Μάγδας κατήγγειλε πολλούς χρηματισμούς μελών της Ολυμπιακής επιτροπής στην κυβέρνηση, αλλά δεν ανέφερε πόσους χρηματισμούς του εαυτού της από την επιτροπή είχε ομολογήσει η φίλη της μόλις την προηγούμενη ημέρα.

iii. Η αδερφή του Αντώνη κατήγγειλε πολλούς χρηματισμούς μελών της Ολυμπιακής επιτροπής στην κυβέρνηση, αλλά δεν ανέφερε για πόσους χρηματισμούς του εαυτού της από την επιτροπή είχε ομολογήσει η φίλη της μόλις την προηγούμενη ημέρα.

iv. Ο αδερφός της Μάγδας κατήγγειλε πολλούς χρηματισμούς μελών της Ολυμπιακής επιτροπής στην κυβέρνηση, αλλά δεν ανέφερε για πόσους χρηματισμούς του εαυτού της από την επιτροπή είχε ομολογήσει η φίλη της μόλις την προηγούμενη ημέρα.

11. i. Η αδερφή του Νίκου αποκάλυψε αρκετούς χρηματισμούς στο συμβούλιο ιατρών, αλλά δεν ήταν προφανές ακόμα πόσους χρηματισμούς του εαυτού του στο συμβούλιο διαπραγματευόταν ο πρόεδρος του συμβουλίου εδώ και πολλά χρόνια.

ii. Ο αδερφός της Έλλης αποκάλυψε αρκετούς χρηματισμούς στο συμβούλιο ιατρών, αλλά δεν ήταν προφανές ακόμα πόσους χρηματισμούς του εαυτού του στο συμβούλιο διαπραγματευόταν ο πρόεδρος του συμβουλίου εδώ και πολλά χρόνια.

iii. Η αδερφή του Νίκου αποκάλυψε αρκετούς χρηματισμούς στο συμβούλιο ιατρών, αλλά δεν ήταν προφανές ακόμα για πόσους χρηματισμούς του εαυτού του στο συμβούλιο διαπραγματευόταν ο πρόεδρος του συμβουλίου εδώ και πολλά χρόνια.

iv. Ο αδερφός της Έλλης αποκάλυψε αρκετούς χρηματισμούς στο συμβούλιο ιατρών, αλλά δεν ήταν προφανές ακόμα για πόσους χρηματισμούς του εαυτού του στο συμβούλιο διαπραγματευόταν ο πρόεδρος του συμβουλίου εδώ και πολλά χρόνια.

12. i. Η αδερφή του Χάρη παρουσίασε πολλούς πίνακες στην έκθεση, αλλά δεν μπορούσε κανείς να καταλάβει ποιον πίνακα του εαυτού της από την έκθεση χάρηκε το μοντέλο

σιωπηρά.

ii. Ο αδερφός της Χαράς παρουσίασε πολλούς πίνακες στην έκθεση, αλλά δεν μπορούσε κανείς να καταλάβει ποιον πίνακα του εαυτού της από την έκθεση χάρηκε το μοντέλο σιωπηρά.

iii. Η αδερφή του Χάρη παρουσίασε πολλούς πίνακες στην έκθεση, αλλά δεν μπορούσε κανείς να καταλάβει για ποιον πίνακα του εαυτού της από την έκθεση χάρηκε το μοντέλο σιωπηρά.

iv. Ο αδερφός της Χαράς παρουσίασε πολλούς πίνακες στην έκθεση, αλλά δεν μπορούσε κανείς να καταλάβει για ποιον πίνακα του εαυτού της από την έκθεση χάρηκε το μοντέλο σιωπηρά.

13. i. Ο ανηψιός του Οδυσσέα μας διάβαζε κάποιες παραγράφους από το διαφημιστικό φυλλάδιο, και όλοι αναρωτιόμασταν πόσες παραγράφους για τον εαυτό της στο φυλλάδιο είχε γράψει η δήμαρχος για προώθηση του κόμματός της.

ii. Η ανηψιά της Αργυρώς μας διάβαζε κάποιες παραγράφους από το διαφημιστικό φυλλάδιο, και όλοι αναρωτιόμασταν πόσες παραγράφους για τον εαυτό της στο φυλλάδιο είχε γράψει η δήμαρχος για προώθηση του κόμματός της.

iii. Ο ανηψιός του Οδυσσέα μας διάβαζε κάποιες παραγράφους από το διαφημιστικό φυλλάδιο, και όλοι αναρωτιόμασταν πόσες παραγράφους για τον εαυτό της στο φυλλάδιο είχε γράψει η δήμαρχος για προώθηση του κόμματός της.

iv. Η ανηψιά της Αργυρώς μας διάβαζε κάποιες παραγράφους από το διαφημιστικό φυλλάδιο, και όλοι αναρωτιόμασταν πόσες παραγράφους για τον εαυτό της στο φυλλάδιο είχε γράψει η δήμαρχος για προώθηση του κόμματός της.

14. i. Ο ανηψιός της Ισμήνης δημοσίευσε κάποιους σχολιασμούς στη σχολική εφημερίδα, αλλά κανείς δεν ξέρει πόσους σχολιασμούς για τον εαυτό του στην εφημερίδα διάβασε ο πατέρας του αργότερα και συγκινήθηκε τόσο πολύ.

ii. Η ανηψιά του Ιάσονα δημοσίευσε κάποιους σχολιασμούς στη σχολική εφημερίδα, αλλά κανείς δεν ξέρει πόσους σχολιασμούς για τον εαυτό του στην εφημερίδα διάβασε ο πατέρας της αργότερα και συγκινήθηκε τόσο πολύ.

iii. Ο ανηψιός της Ισμήνης δημοσίευσε κάποιους σχολιασμούς στη σχολική εφημερίδα, αλλά κανείς δεν ξέρει για πόσους σχολιασμούς για τον εαυτό του στην εφημερίδα διάβασε ο πατέρας του αργότερα και συγκινήθηκε τόσο πολύ.

iv. Η ανηψιά του Ιάσονα δημοσίευσε κάποιους σχολιασμούς στη σχολική εφημερίδα, αλλά κανείς δεν ξέρει για πόσους σχολιασμούς για τον εαυτό του στην εφημερίδα διάβασε ο πατέρας της αργότερα και συγκινήθηκε τόσο πολύ.

15. i. Ο ανηψιός της Σοφίας συνέθεσε κάποιους στίχους για τη σχολική γιορτή, αλλά κανείς δεν μπορούσε να φανταστεί πόσους στίχους για τον εαυτό του στη γιορτή θα τραγουδούσε ο διευθυντής γεμάτος χαρά.

    ii. Η ανηψιά του Θανάση συνέθεσε κάποιους στίχους για τη σχολική γιορτή, αλλά κανείς δεν μπορούσε να φανταστεί πόσους στίχους για τον εαυτό του στη γιορτή θα τραγουδούσε ο διευθυντής γεμάτος χαρά.

    iii. Ο ανηψιός της Σοφίας συνέθεσε κάποιους στίχους για τη σχολική γιορτή, αλλά κανείς δεν μπορούσε να φανταστεί για πόσους στίχους για τον εαυτό του στη γιορτή θα τραγουδούσε ο διευθυντής γεμάτος χαρά.

    iv. Η ανηψιά του Θανάση συνέθεσε κάποιους στίχους για τη σχολική γιορτή, αλλά κανείς δεν μπορούσε να φανταστεί για πόσους στίχους για τον εαυτό του στη γιορτή θα τραγουδούσε ο διευθυντής γεμάτος χαρά.

16. i. Η ανηψιά του Θωμά συνέγραψε κάποιες παραγράφους στο ίντερνετ, αλλά κανείς δεν μπορούσε να φανταστεί πόσες παραγράφους για τον εαυτό του από την ιστοσελίδα θα διάβαζε η γυναίκα του αργότερα και θα θύμωνε.

    ii. Ο ανηψιός της Ζωής συνέγραψε κάποιες παραγράφους στο ίντερνετ, αλλά κανείς δεν μπορούσε να φανταστεί πόσες παραγράφους για τον εαυτό του από την ιστοσελίδα θα διάβαζε η γυναίκα του αργότερα και θα θύμωνε.

    iii. Η ανηψιά του Θωμά συνέγραψε κάποιες παραγράφους στο ίντερνετ, αλλά κανείς δεν μπορούσε να φανταστεί για πόσες παραγράφους για τον εαυτό του από την ιστοσελίδα θα διαβαζε η γυναίκα του αργότερα και θα θύμωνε.

    iv. Ο ανηψιός της Ζωής συνέγραψε κάποιες παραγράφους στο ίντερνετ, αλλά κανείς δεν μπορούσε να φανταστεί για πόσες παραγράφους για τον εαυτό του από την ιστοσελίδα θα διάβαζε η γυναίκα του αργότερα και θα θύμωνε.

17. i. Ο θείος της Εύας δικαιολόγησε ορισμένους χαρακτηρισμούς ανθρώπων στη δημόσια επιτροπή, αλλά κανείς δεν ξέρει ποιο χαρακτηρισμό του εαυτού της από την εφημερίδα πληροφορήθηκε η μητέρα της το ίδιο πρωί με αγανάκτηση.

    ii. Η θεία του Φώτη δικαιολόγησε ορισμένους χαρακτηρισμούς ανθρώπων στη δημόσια επιτροπή, αλλά κανείς δεν ξέρει ποιο χαρακτηρισμό του εαυτού της από την εφημερίδα πληροφορήθηκε η μητέρα της το ίδιο πρωί με αγανάκτηση.

    iii. Ο θείος της Εύας δικαιολόγησε ορισμένους χαρακτηρισμούς ανθρώπων στη δημόσια επιτροπή, αλλά κανείς δεν ξέρει για ποιο χαρακτηρισμό του εαυτού της από την εφημερίδα πληροφορήθηκε η μητέρα της το ίδιο πρωί με αγανάκτηση.

    iv. Η θεία του Φώτη δικαιολόγησε ορισμένους χαρακτηρισμούς ανθρώπων στη δημόσια επιτροπή, αλλά κανείς δεν ξέρει για ποιο χαρακτηρισμό του εαυτού της από την εφημερίδα πληροφορήθηκε η μητέρα της το ίδιο πρωί με αγανάκτηση.

18. i. Ο θείος της Παναγιώτας παρακολουθούσε κάποιους σχολιασμούς στο δελτίο ειδήσεων, και ήταν προφανές πόσους σχολιασμούς για τον εαυτό της στις ειδήσεις θα φοβόταν η βουλευτής έχοντας προκαλέσει τους δημοσιογράφους.

    ii. Η θεία του Παναγιώτη παρακολουθούσε κάποιους σχολιασμούς στο δελτίο ειδήσεων, και ήταν προφανές πόσους σχολιασμούς για τον εαυτό της στις ειδήσεις θα φοβόταν η βουλευτής έχοντας προκαλέσει τους δημοσιογράφους.

    iii. Ο θείος της Παναγιώτας παρακολουθούσε κάποιους σχολιασμούς στο δελτίο ειδήσεων, και ήταν προφανές για πόσους σχολιασμούς για τον εαυτό της στις ειδήσεις θα φοβόταν η βουλευτής έχοντας προκαλέσει τους δημοσιογράφους.

    iv. Η θεία του Παναγιώτη παρακολουθούσε κάποιους σχολιασμούς στο δελτίο ειδήσεων, και ήταν προφανές για πόσους σχολιασμούς για τον εαυτό της στις ειδήσεις θα φοβόταν η βουλευτής έχοντας προκαλέσει τους δημοσιογράφους.

19. i. Η θεία του Αριστείδη βρήκε κάποιους σχολικούς επαίνους σε ένα συρτάρι, αλλά δε θυμάμαι πόσους επαίνους του εαυτού της από το συρτάρι ερεύνησε η ανηψιά του αργότερα όταν είχαν φύγει όλοι.

    ii. Ο θείος της Βασιλείας βρήκε κάποιους σχολικούς επαίνους σε ένα συρτάρι, αλλά δε θυμάμαι πόσους επαίνους του εαυτού της από το συρτάρι ερεύνησε η ανηψιά του αργότερα όταν είχαν φύγει όλοι.

    iii. Η θεία του Αριστείδη βρήκε κάποιους σχολικούς επαίνους σε ένα συρτάρι, αλλά δε θυμάμαι για πόσους επαίνους του εαυτού της από το συρτάρι ερεύνησε η ανηψιά του αργότερα όταν είχαν φύγει όλοι.

    iv. Ο θείος της Βασιλείας βρήκε κάποιους σχολικούς επαίνους σε ένα συρτάρι, αλλά δε θυμάμαι για πόσους επαίνους του εαυτού της από το συρτάρι ερεύνησε η ανηψιά του αργότερα όταν είχαν φύγει όλοι.

20. i. Ο θείος της Βάσως ευχαρίστησε κάποιους ευεργέτες της εταιρείας στη δημόσια ομιλία, αλλά κανείς δεν θα μπορούσε να φανταστεί πόσους ευεργέτες του εαυτού του από το παρελθόν πληροφορήθηκε ο διευθυντής της εταιρείας μετά την ομιλία.

    ii. Η θεία του Τάκη ευχαρίστησε κάποιους ευεργέτες της εταιρείας στη δημόσια ομιλία, αλλά κανείς δεν θα μπορούσε να φανταστεί πόσους ευεργέτες του εαυτού του από το παρελθόν πληροφορήθηκε ο διευθυντής της εταιρείας μετά την ομιλία.

    iii. Ο θείος της Βάσως ευχαρίστησε κάποιους ευεργέτες της εταιρείας στη δημόσια ομιλία, αλλά κανείς δεν θα μπορούσε να φανταστεί για πόσους ευεργέτες του εαυτού του από το παρελθόν πληροφορήθηκε ο διευθυντής της εταιρείας μετά την ομιλία.

iv. Η θεία του Τάκη ευχαρίστησε κάποιους ευεργέτες της εταιρείας στη δημόσια ομιλία, αλλά κανείς δεν θα μπορούσε να φανταστεί για πόσους ευεργέτες του εαυτού του από το παρελθόν πληροφορήθηκε ο διευθυντής της εταιρείας μετά την ομιλία.

21. i. Η γυναίκα του Πάρη δέχτηκε πολλούς υποστηρικτές στις τοπικές εκλογές, αλλά δεν ήταν ξεκάθαρο ποιον υποστηρικτή του εαυτού της στις εκλογές ντράπηκε η πολιτική αντίπαλός της τόσο υπέρμετρα.

    ii. Ο άνδρας της Άννας δέχτηκε πολλούς υποστηρικτές στις τοπικές εκλογές, αλλά δεν ήταν ξεκάθαρο ποιον υποστηρικτή του εαυτού της στις εκλογές ντράπηκε η πολιτική αντίπαλός της τόσο υπέρμετρα.

    iii. Η γυναίκα του Πάρη δέχτηκε πολλούς υποστηρικτές στις τοπικές εκλογές, αλλά δεν ήταν ξεκάθαρο για ποιον υποστηρικτή του εαυτού της στις εκλογές ντράπηκε η πολιτική αντίπαλός της τόσο υπέρμετρα.

    iv. Ο άνδρας της Άννας δέχτηκε πολλούς υποστηρικτές στις τοπικές εκλογές, αλλά δεν ήταν ξεκάθαρο για ποιον υποστηρικτή του εαυτού της στις εκλογές ντράπηκε η πολιτική αντίπαλός της τόσο υπέρμετρα.

22. i. Η κοπέλα του Δημήτρη βρήκε κάποιους στίχους σε ένα ρομαντικό ραβασάκι, αλλά κανείς δε θυμάται ποιους στίχους για τον εαυτό της στο ραβασάκι διάβασε η αδερφή της και ξέσπασε σε κλάματα.

    ii. Ο νεαρός της Δήμητρας βρήκε κάποιους στίχους σε ένα ρομαντικό ραβασάκι, αλλά κανείς δε θυμάται ποιους στίχους για τον εαυτό της στο ραβασάκι διάβασε η αδερφή της και ξέσπασε σε κλάματα.

    iii. Η κοπέλα του Δημήτρη βρήκε κάποιους στίχους σε ένα ρομαντικό ραβασάκι, αλλά κανείς δε θυμάται για ποιους στίχους για τον εαυτό της στο ραβασάκι διάβασε η αδερφή της και ξέσπασε σε κλάματα.

    iv. Ο νεαρός της Δήμητρας βρήκε κάποιους στίχους σε ένα ρομαντικό ραβασάκι, αλλά κανείς δε θυμάται για ποιους στίχους για τον εαυτό της στο ραβασάκι διάβασε η αδερφή της και ξέσπασε σε κλάματα.

23. i. Η κοπέλα του Λευτέρη αντιμετώπισε πολλούς κριτές στο τουρνουά ενόργανης γυμναστικής, αλλά δεν ήταν ξεκάθαρο πόσους κριτές του εαυτού της από το τουρνουά φοβόταν η φίλη της από την ίδια ομάδα.

    ii. Ο νεαρός της Μαρίνας αντιμετώπισε πολλούς κριτές στο τουρνουά ενόργανης γυμναστικής, αλλά δεν ήταν ξεκάθαρο πόσους κριτές του εαυτού της από το τουρνουά φοβόταν η φίλη της από την ίδια ομάδα.

iii. Η κοπέλα του Λευτέρη αντιμετώπισε πολλούς κριτές στο τουρνουά ενόργανης γυμναστικής, αλλά δεν ήταν ξεκάθαρο για πόσους κριτές του εαυτού της από το τουρνουά φοβόταν η φίλη της από την ίδια ομάδα.

iv. Ο νεαρός της Μαρίνας αντιμετώπισε πολλούς κριτές στο τουρνουά ενόργανης γυμναστικής, αλλά δεν ήταν ξεκάθαρο για πόσους κριτές του εαυτού της από το τουρνουά φοβόταν η φίλη της από την ίδια ομάδα.

24. i. Η κοπέλα του Ματθαίου μάζεψε αρκετούς υποστηρικτές στις σχολικές εκλογές, αλλά δεν έχω υπόψιν μου πόσους υποστηρικτές του εαυτού του στις εκλογές συναναστρέφεται ο πρώην της ιδιαίτερα συχνά.

ii. Ο νεαρός της Δανάης μάζεψε αρκετούς υποστηρικτές στις σχολικές εκλογές, αλλά δεν έχω υπόψιν μου πόσους υποστηρικτές του εαυτού του στις εκλογές συναναστρέφεται ο πρώην της ιδιαίτερα συχνά.

iii. Η κοπέλα του Ματθαίου μάζεψε αρκετούς υποστηρικτές στις σχολικές εκλογές, αλλά δεν έχω υπόψιν μου με πόσους υποστηρικτές του εαυτού του στις εκλογές συναναστρέφεται ο πρώην της ιδιαίτερα συχνά.

iv. Ο νεαρός της Δανάης μάζεψε αρκετούς υποστηρικτές στις σχολικές εκλογές, αλλά δεν έχω υπόψιν μου γμε πόσους υποστηρικτές του εαυτού του στις εκλογές συναναστρέφεται ο πρώην της ιδιαίτερα συχνά.

25. i. Ο αδερφός της Ρωξάνης αναγνώρισε την ύπαρξη λίγων μιμητών στο χώρο της τηλεόρασης, χωρίς να μας αποκαλύψει ποιον μιμητή του εαυτού του στην τηλεόραση είχε προειδοποιήσει ο διάσημος παρουσιαστής του βραδινού δελτίου.

ii. Η αδερφή του Ρωμύλου αναγνώρισε την ύπαρξη λίγων μιμητών στο χώρο της τηλεόρασης, χωρίς να μας αποκαλύψει ποιον μιμητή του εαυτού του στην τηλεόραση είχε προειδοποιήσει ο διάσημος παρουσιαστής του βραδινού δελτίου.

iii. Ο αδερφός της Ρωξάνης αναγνώρισε την ύπαρξη λίγων μιμητών στο χώρο της τηλεόρασης, χωρίς να μας αποκαλύψει για ποιον μιμητή του εαυτού του στην τηλεόραση είχε προειδοποιήσει ο διάσημος παρουσιαστής του βραδινού δελτίου.

iv. Η αδερφή του Ρωμύλου αναγνώρισε την ύπαρξη λίγων μιμητών στο χώρο της τηλεόρασης, χωρίς να μας αποκαλύψει για ποιον μιμητή του εαυτού του στην τηλεόραση είχε προειδοποιήσει ο διάσημος παρουσιαστής του βραδινού δελτίου.

26. i. Η φίλη του Δημοσθένη αντιμετώπισε με ευκολία πολλούς εχθρούς του κόμματός της στη Βουλή, αλλά όλοι κατάλαβαν ποιον εχθρό του εαυτού της στη Βουλή σώπασε η συνάδελφός της που μπήκε λίγο αργότερα στο κοινοβούλιο.

ii. Ο φίλος της Περσεφόνης αντιμετώπισε με ευκολία πολλούς εχθρούς του κόμματός της στη Βουλή, αλλά όλοι κατάλαβαν ποιον εχθρό του εαυτού της στη Βουλή σώπασε η συνάδελφός του που μπήκε λίγο αργότερα στο κοινοβούλιο.

iii. Η φίλη του Δημοσθένη αντιμετώπισε με ευκολία πολλούς εχθρούς του κόμματός της στη Βουλή, αλλά όλοι κατάλαβαν για ποιον εχθρό του εαυτού της στη Βουλή σώπασε η συνάδελφός της που μπήκε λίγο αργότερα στο κοινοβούλιο.

iv. Ο φίλος της Περσεφόνης αντιμετώπισε με ευκολία πολλούς εχθρούς του κόμματός της στη Βουλή, αλλά όλοι κατάλαβαν για ποιον εχθρό του εαυτού της στη Βουλή σώπασε η συνάδελφός του που μπήκε λίγο αργότερα στο κοινοβούλιο.

27. i. Η φίλη του Σωκράτη διάβασε κάποιες παραγράφους για τους συμμαθητές στη σχολική επετηρίδα, αλλά δεν ήταν ξεκάθαρο ποια παράγραφο για τον εαυτό του στην επετηρίδα πληροφορήθηκε ο απουσιολόγος της τάξης και εκνευρίστηκε.

ii. Ο φίλος της Αγάπης διάβασε κάποιες παραγράφους για τους συμμαθητές στη σχολική επετηρίδα, αλλά δεν ήταν ξεκάθαρο ποια παράγραφο για τον εαυτό του στην επετηρίδα πληροφορήθηκε ο απουσιολόγος της τάξης και εκνευρίστηκε.

iii. Η φίλη του Σωκράτη διάβασε κάποιες παραγράφους για τους συμμαθητές στη σχολική επετηρίδα, αλλά δεν ήταν ξεκάθαρο για ποια παράγραφο για τον εαυτό στην επετηρίδα πληροφορήθηκε ο απουσιολόγος της τάξης και εκνευρίστηκε.

iv. Ο φίλος της Αγάπης διάβασε κάποιες παραγράφους για τους συμμαθητές στη σχολική επετηρίδα, αλλά δεν ήταν ξεκάθαρο για ποια παράγραφο για τον εαυτό στην επετηρίδα πληροφορήθηκε ο απουσιολόγος της τάξης και εκνευρίστηκε.

28. i. Η νονά του Αναστάση ανέβασε κάποιους πίνακές της στο ίντερνετ προς πώληση, αλλά δεν θυμόμασταν πόσους πίνακες του εαυτού της στο ίντερνετ ευχαριστήθηκε η αδερφή της τόσο πολύ όταν τους είδε.

ii. Ο νονός της Περσεφόνης ανέβασε κάποιους πίνακές του στο ίντερνετ προς πώληση, αλλά δεν θυμόμασταν πόσους πίνακες του εαυτού της στο ίντερνετ ευχαριστήθηκε η αδερφή της τόσο πολύ όταν τους είδε.

iii. Η νονά του Αναστάση ανέβασε κάποιους πίνακές της στο ίντερνετ προς πώληση, αλλά δεν θυμόμασταν πόσους πίνακες του εαυτού της στο ίντερνετ ευχαριστήθηκε η αδερφή της τόσο πολύ όταν τους είδε.

iv. Ο νονός της Περσεφόνης ανέβασε κάποιους πίνακές του στο ίντερνετ προς πώληση, αλλά δεν θυμόμασταν πόσους πίνακες του εαυτού της στο ίντερνετ ευχαριστήθηκε η αδερφή της τόσο πολύ όταν τους είδε.

29. i. Ο αδερφός της Ράνιας αναγνώρισε πολλούς ισάξιους μέσα στο διοικητικό συμβούλιο, αλλά δεν αποκάλυψε ποιον ισάξιο του εαυτού του μέσα στο συμβούλιο άκουγε ο

πρόεδρος με προσοχή.

ii. Η αδερφή του Σεραφείμ αναγνώρισε πολλούς ισάξιους μέσα στο διοικητικό συμβούλιο, αλλά δεν αποκάλυψε ποιον ισάξιο του εαυτού του μέσα στο συμβούλιο άκουγε ο πρόεδρος με προσοχή.

iii. Ο αδερφός της Ράνιας αναγνώρισε πολλούς ισάξιους μέσα στο διοικητικό συμβούλιο, αλλά δεν αποκάλυψε για ποιον ισάξιο του εαυτού του μέσα στο συμβούλιο άκουγε ο πρόεδρος με προσοχή.

iv. Η αδερφή του Σεραφείμ αναγνώρισε πολλούς ισάξιους μέσα στο διοικητικό συμβούλιο, αλλά δεν αποκάλυψε για ποιον ισάξιο του εαυτού του μέσα στο συμβούλιο άκουγε ο πρόεδρος με προσοχή.

30. i. Η ιατρός του Χρύσανθου καυτηρίασε τους επικριτές των γιατρών του νοσοκομείου, αλλά δεν κατάλαβε κανείς ποιον επικριτή του εαυτού της μέσα στο νοσοκομείο ντράπηκε η συνάδελφός της στο ιατρικό συμβούλιο.

ii. Ο ιατρός της Χρυσάνθης καυτηρίασε τους επικριτές των γιατρών του νοσοκομείου, αλλά δεν κατάλαβε κανείς ποιον επικριτή του εαυτού της μέσα στο νοσοκομείο ντράπηκε η συνάδελφός της στο ιατρικό συμβούλιο.

iii. Η ιατρός του Χρύσανθου καυτηρίασε τους επικριτές των γιατρών του νοσοκομείου, αλλά δεν κατάλαβε κανείς για ποιον επικριτή του εαυτού της μέσα στο νοσοκομείο ντράπηκε η συνάδελφός της στο ιατρικό συμβούλιο.

iv. Ο ιατρός της Χρυσάνθης καυτηρίασε τους επικριτές των γιατρών του νοσοκομείου, αλλά δεν κατάλαβε κανείς για ποιον επικριτή του εαυτού της μέσα στο νοσοκομείο ντράπηκε η συνάδελφός της στο ιατρικό συμβούλιο.

31. i. Ο γαμπρός της Πολυάννας διαμαρτυρήθηκε για ορισμένους στιγματισμούς των μελών της οργάνωσης, αλλά κανείς δεν έμεινε να ακούσει πόσους στιγματισμούς του εαυτού του από την οργάνωση καταριόταν ο θείος του όλο το βράδυ.

ii. Η νύφη του Τηλέμαχου διαμαρτυρήθηκε για ορισμένους στιγματισμούς των μελών της οργάνωσης, αλλά κανείς δεν έμεινε να ακούσει πόσους στιγματισμούς του εαυτού του από την οργάνωση καταριόταν ο θείος του όλο το βράδυ.

iii. Ο γαμπρός της Πολυάννας διαμαρτυρήθηκε για ορισμένους στιγματισμούς των μελών της οργάνωσης, αλλά κανείς δεν έμεινε να ακούσει για πόσους στιγματισμούς του εαυτού του από την οργάνωση καταριόταν ο θείος του όλο το βράδυ.

iv. Η νύφη του Τηλέμαχου διαμαρτυρήθηκε για ορισμένους στιγματισμούς των μελών της οργάνωσης, αλλά κανείς δεν έμεινε να ακούσει για πόσους στιγματισμούς του εαυτού του από την οργάνωση καταριόταν ο θείος του όλο το βράδυ.

32. i. Ο γιόκας της Πετρούλας συνέγραψε μερικούς ύμνους για διάφορα άτομα στο σχολείο, όμως δεν ήταν εμφανές πόσους ύμνους για τον εαυτό της από τα διάφορα παιδιά ευχαριστήθηκε η δασκάλα του.

    ii. Η κορούλα του Προδρόμου συνέγραψε μερικούς ύμνους για διάφορα άτομα στο σχολείο, όμως δεν ήταν εμφανές πόσους ύμνους για τον εαυτό της από τα διάφορα παιδιά ευχαριστήθηκε η δασκάλα της.

    iii. Ο γιόκας της Πετρούλας συνέγραψε μερικούς ύμνους για διάφορα άτομα στο σχολείο, όμως δεν ήταν εμφανές για πόσους ύμνους για τον εαυτό της από τα διάφορα παιδιά ευχαριστήθηκε η δασκάλα του.

    iv. Η κορούλα του Προδρόμου συνέγραψε μερικούς ύμνους για διάφορα άτομα στο σχολείο, όμως δεν ήταν εμφανές για πόσους ύμνους για τον εαυτό της από τα διάφορα παιδιά ευχαριστήθηκε η δασκάλα της.

33. i. Η κόρη του Αδαμάντιου δέχτηκε πολλούς επαίνους στην πρεμιέρα της παράστασης, όμως δε θυμάμαι πόσους επαίνους για τον εαυτό της στην πρεμιέρα ευχαριστήθηκε η πρωταγωνίστρια της παράστασης εξίσου.

    ii. Ο γιος της Αδαμαντίας δέχτηκε πολλούς επαίνους στην πρεμιέρα της παράστασης, όμως δε θυμάμαι πόσους επαίνους για τον εαυτό της στην πρεμιέρα ευχαριστήθηκε η πρωταγωνίστρια της παράστασης εξίσου.

    iii. Η κόρη του Αδαμάντιου δέχτηκε πολλούς επαίνους στην πρεμιέρα της παράστασης, όμως δε θυμάμαι για πόσους επαίνους για τον εαυτό της στην πρεμιέρα ευχαριστήθηκε η πρωταγωνίστρια της παράστασης εξίσου.

    iv. Ο γιος της Αδαμαντίας δέχτηκε πολλούς επαίνους στην πρεμιέρα της παράστασης, όμως δε θυμάμαι για πόσους επαίνους για τον εαυτό της στην πρεμιέρα ευχαριστήθηκε η πρωταγωνίστρια της παράστασης εξίσου.

34. i. Ο δικηγόρος της Σταματίνας αναφέρθηκε εκτενώς σε κάποιους αντίδικους της προηγούμενης εβδομάδας, και ήταν φανερό ποιον αντίδικο του εαυτού του από την περασμένη εβδομάδα αγανάκτησε ο συνάδελφός του περισσότερο από όλους.

    ii. Η δικηγόρος του Σταμάτη αναφέρθηκε εκτενώς σε κάποιους αντίδικους της προηγούμενης εβδομάδας, και ήταν φανερό ποιον αντίδικο του εαυτού του από την περασμένη εβδομάδα αγανάκτησε ο συνάδελφός του περισσότερο από όλους.

    iii. Ο δικηγόρος της Σταματίνας αναφέρθηκε εκτενώς σε κάποιους αντίδικους της προηγούμενης εβδομάδας, και ήταν φανερό με ποιον αντίδικο του εαυτού τουα πό την περασμένη εβδομάδα αγανάκτησε ο συνάδελφός του περισσότερο από όλους.

    iv. Η δικηγόρος του Σταμάτη αναφέρθηκε εκτενώς σε κάποιους αντίδικους της προηγούμενης εβδομάδας, και ήταν φανερό με ποιον αντίδικο του εαυτού του από την περασμένη εβδομάδα αγανάκτησε ο συνάδελφός του περισσότερο από όλους.

35. i. Ο ξάδερφος της Ιωάννας αψηφούσε τους αντιπάλους της ομάδας του στη σχολική αυλή, αλλά ήταν προφανές ποιον αντίπαλο του εαυτού του στη σχολική αυλή φοβόταν ο μικρότερος αδερφός του.

 ii. Η ξαδέρφη του Παντελή αψηφούσε τους αντιπάλους της ομάδας της στη σχολική αυλή, αλλά ήταν προφανές ποιον αντίπαλο του εαυτού του στη σχολική αυλή φοβόταν ο μικρότερος αδερφός του.

 iii. Ο ξάδερφος της Ιωάννας αψηφούσε τους αντιπάλους της ομάδας του στη σχολική αυλή, αλλά ήταν προφανές για ποιον αντίπαλο του εαυτού του στη σχολική αυλή φοβόταν ο μικρότερος αδερφός του.

 iv. Η ξαδέρφη του Παντελή αψηφούσε τους αντιπάλους της ομάδας της στη σχολική αυλή, αλλά ήταν προφανές για ποιον αντίπαλο του εαυτού του στη σχολική αυλή φοβόταν ο μικρότερος αδερφός του.

36. i. Ο μπαμπάς της Μυρτώς επαίνεσε ορισμένους βιογράφους του καλλιτεχνικού κύκλου στην εκδήλωση, αλλά δε θυμάμαι ποιον βιογράφο του εαυτού του στην εκδήλωση αγανάκτησε ο διάσημος ζωγράφος στο τέλος.

 ii. Η μαμά του Παύλου επαίνεσε ορισμένους βιογράφους του καλλιτεχνικού κύκλου στην εκδήλωση, αλλά δε θυμάμαι ποιον βιογράφο του εαυτού του στην εκδήλωση αγανάκτησε ο διάσημος ζωγράφος στο τέλος.

 iii. Ο μπαμπάς της Μυρτώς επαίνεσε ορισμένους βιογράφους του καλλιτεχνικού κύκλου στην εκδήλωση, αλλά δε θυμάμαι με ποιον βιογράφο του εαυτού του στην εκδήλωση αγανάκτησε ο διάσημος ζωγράφος στο τέλος.

 iv. Η μαμά του Παύλου επαίνεσε ορισμένους βιογράφους του καλλιτεχνικού κύκλου στην εκδήλωση, αλλά δε θυμάμαι με ποιον βιογράφο του εαυτού του στην εκδήλωση αγανάκτησε ο διάσημος ζωγράφος στο τέλος.

37. i. Ο παππούς της Φανής επαινούσε ορισμένους αντιπάλους της παλιάς ομάδας του στο ποδόσφαιρο, αλλά δεν θυμάμαι πόσους αντιπάλους του εαυτού του στο μπάσκετ έτρεμε ο αθλητής αδερφός του όσο έπαιζε.

 ii. Η γιαγιά του Ζήση επαινούσε ορισμένους αντιπάλους της παλιάς ομάδας της στο ποδόσφαιρο, αλλά δεν θυμάμαι πόσους αντιπάλους του εαυτού του στο μπάσκετ έτρεμε ο αθλητής αδερφός της όσο έπαιζε.

 iii. Ο παππούς της Φανής επαινούσε ορισμένους αντιπάλους της παλιάς ομάδας του στο ποδόσφαιρο, αλλά δεν θυμάμαι για πόσους αντιπάλους του εαυτού του στο μπάσκετ έτρεμε ο αθλητής αδερφός του όσο έπαιζε.

iv. Η γιαγιά του Ζήση επαινούσε ορισμένους αντιπάλους της παλιάς ομάδας της στο ποδόσφαιρο, αλλά δεν θυμάμαι για πόσους αντιπάλους του εαυτού του στο μπάσκετ έτρεμε ο αθλητής αδερφός της όσο έπαιζε.

38. i. Ο καθηγητής της Αριάδνης ανέδειξε κάποιους καθρέφτες της ψυχής μας από την παιδική μας ηλικία, αλλά δεν αποκάλυψε ποιον καθρέφτη του εαυτού του από την παιδική του ηλικία γνώριζε ο ποιητής .

ii. Η καθηγήτρια του Αποστόλη ανέδειξε κάποιους καθρέφτες της ψυχής μας από την παιδική μας ηλικία, αλλά δεν αποκάλυψε ποιον καθρέφτη του εαυτού του από την παιδική του ηλικία γνώριζε ο ποιητής .

iii. Ο καθηγητής της Αριάδνης ανέδειξε κάποιους καθρέφτες της ψυχής μας από την παιδική μας ηλικία, αλλά δεν αποκάλυψε ποιον καθρέφτη του εαυτού του από την παιδική του ηλικία γνώριζε ο ποιητής .

iv. Η καθηγήτρια του Αποστόλη ανέδειξε κάποιους καθρέφτες της ψυχής μας από την παιδική μας ηλικία, αλλά δεν αποκάλυψε ποιον καθρέφτη του εαυτού του από την παιδική του ηλικία γνώριζε ο ποιητής .

39. i. Ο νονός της Αργυρώς σχολίασε ορισμένους αποκλεισμούς επιστημόνων από αποστολές του περασμένου αιώνα, και θυμάμαι ακριβώς πόσους αποκλεισμούς του εαυτού της από αποστολές συζήτησε η παλιά εξερευνήτρια αργότερα.

ii. Η νονά του Περικλή σχολίασε ορισμένους αποκλεισμούς επιστημόνων από αποστολές του περασμένου αιώνα, και θυμάμαι ακριβώς πόσους αποκλεισμούς του εαυτού της από αποστολές συζήτησε η παλιά εξερευνήτρια αργότερα.

iii. Ο νονός της Αργυρώς σχολίασε ορισμένους αποκλεισμούς επιστημόνων από αποστολές του περασμένου αιώνα, και θυμάμαι ακριβώς πόσους αποκλεισμούς του εαυτού της από αποστολές συζήτησε η παλιά εξερευνήτρια αργότερα.

iv. Η νονά του Περικλή σχολίασε ορισμένους αποκλεισμούς επιστημόνων από αποστολές του περασμένου αιώνα, και θυμάμαι ακριβώς πόσους αποκλεισμούς του εαυτού της από αποστολές συζήτησε η παλιά εξερευνήτρια αργότερα.

40. i. Η αδερφή του Προμυθέα εξύμνησε πολλούς υποστηρικτές προσφύγων στον καταυλισμό των συνόρων, αλλά δε θυμάμαι πόσους υποστηρικτές του εαυτού της στα σύνορα συναναστρεφόταν η ταλαίπωρη πρόσφυγας στον καταυλισμό.

ii. Ο αδερφός της Ουρανίας εξύμνησε πολλούς υποστηρικτές προσφύγων στον καταυλισμό των συνόρων, αλλά δε θυμάμαι πόσους υποστηρικτές του εαυτού της στα σύνορα συναναστρεφόταν η ταλαίπωρη πρόσφυγας στον καταυλισμό.

iii. Η αδερφή του Προμυθέα εξύμνησε πολλούς υποστηρικτές προσφύγων στον καταυλισμό των συνόρων, αλλά δε θυμάμαι πόσους υποστηρικτές του εαυτού της στα σύνορα συναναστρεφόταν η ταλαίπωρη πρόσφυγας στον καταυλισμό.

iv. Ο αδερφός της Ουρανίας εξύμνησε πολλούς υποστηρικτές προσφύγων στον καταυλισμό των συνόρων, αλλά δε θυμάμαι πόσους υποστηρικτές του εαυτού της στα σύνορα συναναστρεφόταν η ταλαίπωρη πρόσφυγας στον καταυλισμό.

Fillers

1. Η κυβέρνηση στις τελευταίες συναντήσεις της σπρώχνει για αξιοποίηση της δημόσιας ακίνητης περιουσίας, αλλά δεν είναι σίγουρο ακόμα ποιες χώρες θα ήθελαν να επενδύσουν σε αυτήν.

2. Μια γνωστή τουριστική εταιρεία ανήγγειλε ότι έχει δημιουργήσει ένα νέο είδος υποδημάτων που σε καθοδηγούν με δονήσεις στον προορισμό σου, και στη συνέντευξη τύπου κάποιοι δημοσιογράφοι παραξενεύτηκαν ενώ άλλοι ενθουσιάστηκαν.

3. Πρόσφατη έρευνα έδειξε ότι τα όνειρα είναι πολύ σημαντικά στην παγίωση των αναμνήσεων, και έτσι εάν αυτά διαταραχθούν υπάρχει περίπτωση να μη λειτουργεί καλά η μνήμη μας την επόμενη ημέρα.

4. Τα πλαϊνά τζάμια δεν προστατεύουν καλά από την υπεριώδη ακτινοβολία του ήλιου, οπότε οι επιστήμονες συνιστούν να φοράμε γυαλιά ηλίου, αντηλιακό και μακριά μανίκια στο αμάξι για καλύτερη προστασία.

5. Η ξαδέρφη του Γρηγόρη μου είπε ότι τα παιδιά που περπατάνε νωρίτερα έχουν πιο γερά κόκαλα, οπότε είχαμε περιέργεια να δούμε πότε θα περπατούσε η μικρότερη κόρη μας.

6. Η Αναστασία εξηγούσε στις φίλες της ότι στην Κροατία έχουν πολλά έθιμα παρόμοια με τα ελληνικά για το Πάσχα, και τους έλεγε για το τσουρέκι και τα κόκκινα αυγά που είχε δει εκεί.

7. Λόγω του ανθυγιεινού τρόπου ζωής της μοντέρνας κοινωνίας η παχυσαρκία αποτελεί πλέον πραγματική επιδημία, ενώ εκτιμάται ότι μέχρι το 2030 θα έχουν προσβληθεί από τη νόσο περισσότεροι από τους μισούς ευρωπαίους πολίτες.

8. Η δασκάλα του Λουκά άργησε να φτάσει στο μάθημα λόγω ενός ατυχήματος στην Εθνική Οδό, αλλά ευτυχώς μαθεύτηκε σύντομα στις ειδήσεις ότι κανείς δεν είχε τραυματιστεί σε αυτό.

9. Ο ξάδερφος της Πόπης πέταξε πολλές εικόνες από το άλμπουμ, αλλά δεν μπόρεσα να δω ποιες εικόνες του εαυτού της από το άλμπουμ άρεσαν στη μητέρα της ιδιαίτερα εμφανώς.

10. Η μητέρα του Ορέστη άκουσε κάποιες φήμες στη σχολική συνάντηση, αλλά ποιος ξέρει ποια φήμη για τον εαυτό του από τη συνάντηση διασκέδασε τον αδερφό του τόσο που γέλασε δυνατά.

11. Η θεία του Αρίστου βρήκε κάποιες παλιές φωτογραφίες στην σκονισμένη αποθήκη, αλλά δε θυμάμαι ποιες φωτογραφίες του εαυτού του στην αποθήκη συγκίνησαν τον πατέρα του τόσο πολύ.

12. Ο ξάδερφος της Κατερίνας είδε πολλά βίντεο στο συνεταιρικό πάρτι, ωστόσο δεν ήταν ιδιαίτερα εμφανές ποιο βίντεο του εαυτού της από το πάρτι συνεπήρε τη γραμματέα του τόσο πολύ.

13. Η βαφτιστήρα του Γιάννη άκουσε κάποιες ηχογραφήσεις στο στούντιο, αλλά δε θα μπορούσε ποτέ να είχε φανταστεί πόσες ηχογραφήσεις του εαυτού του από το στούντιο θα σχολίαζε ο μάνατζέρ της τόσο επικριτικά.

14. Ο γιος της Μαρίκας ανέβασε πολλές φωτογραφίες στο ίντερνετ, αλλά κανείς δε φανταζόταν πόσες φωτογραφίες του εαυτού της από την ιστοσελίδα θα ενοχλούσαν την κοπέλα του σε σημείο να τον χωρίσει.

15. Η κόρη του Ιάκωβου είπε κάποια ενοχλητικά ανέκδοτα στην πλατεία της γειτονιάς, ωστόσο δεν κατάλαβα ποιο ανέκδοτο για τον εαυτό του άκουσε ο γείτονας και εξαγριώθηκε.

16. Η κοπέλα του Φοίβου ανακάλεσε ορισμένες ωραίες αναμνήσεις μια καλοκαιρινή ημέρα, και ήταν εύκολο να καταλάβει κανείς ποιες αναμνήσεις για τον εαυτό της ανακάλεσε η αδερφή της με έναν βαθύ αναστεναγμό.

17. Ο γιος της Αθανασίας έδειξε κάποια σχέδια στους φίλους του, αλλά δεν κατάλαβε κανείς ποιο σχέδιο του εαυτού του είχε δει ο αδερφός της και είχε λατρέψει.

18. Η μικρή Ηλιάνα έμαθε στο σχολείο σήμερα για κάποια νέα έντομα που ανακαλύφθηκαν πρόσφατα, και δεν μπορούσε να σταματήσει να γελάει όταν άκουσε πώς τα ονόμασαν οι επιστήμονες.

19. Η αδερφούλα του Κυριάκου έκανε κάποιες πολύ ωραίες περιγραφές ανθρώπων από το πάρτι, και ήταν προφανές ποια περιγραφή του εαυτού της από το πάρτι απόλαυσε μητέρα της πολύ και γέλασε.

20. Ο βαφτισιμιός της Αλεξάνδρας μας έπαιζε κάποιες ηχογραφήσεις από την πρόβα της χορωδίας, ωστόσο δεν ήταν ξεκάθαρο ποια ηχογράφηση του εαυτού του από την πρόβα διαφήμιζε ο αδερφός του τόσο και κοκορευόταν.

21. Η κόρη της Ρίτας άκουσε ορισμένες φήμες από την τάξη της στο δημοτικό, ωστόσο δε μάθαμε ποια φήμη για τον εαυτό της έμαθε η δίδυμη αδερφή της και παραξενεύτηκε πολύ.

22. Ο μικρός γιος της Ευτυχίας έδειχνε παντού κάποια βίντεο από το τουρνουά ποδηλασίας, αλλά όλοι καταλάβαμε ποιο βίντεο του εαυτού του από το τουρνουά πρόβαλε ο πατέρας του αργότερα στους φίλους του.

23. Η αδερφή του Σωκράτη διηγούνταν κάποιες ιστορίες από την παιδική της ηλικία, αλλά δεν κατάλαβε κανείς ποια ιστορία για τον εαυτό της ντράπηκε η ξαδέρφη της τόσο πολύ που κοκκίνησε.

24. Ο προπάππος της Ευλαμπίας έλεγε πολλές ιστορίες από τον πόλεμο, και εύκολα καταλάβαμε ποια ιστορία του εαυτού του από τον πόλεμο συγκίνησε τόσο τον αδερφό του που δάκρυσε.

25. Ο καθηγητής της φίλης μου εξηγούσε πως ο καιρός στο διάστημα επηρεάζει τον καιρό στη γη, και συζητούσαμε πώς θα μελετούνταν ο καιρός στο διάστημα και τι επιπτώσεις θα έχει αυτό.

26. Ο γιατρός εξηγούσε στη μητέρα μου ότι υπάρχει καλή και κακή χοληστερόλη, όμως αυτή επιμένει να τρώει ανθυγιεινά και να καπνίζει προς μεγάλη μας απογοήτευση.

27. Το κλειδί για ένα πετυχημένο δώρο δεν είναι να αγοράζει κανείς αυτό που αρέσει στον εορτάζοντα, αλλά να αγοράζει αυτό που αρέσει στον εαυτό του για να αποτυπωθεί σε αυτό η προσωπικότητά του.

28. Η μητέρα του Γιωργάκη ήθελε να του πάρει ένα σκυλάκι ως έκπληξη για τη γιορτή του, αλλά δεν μπορούσε να αποφασίσει ποιο σκύλο να σώσει από τη φιλοζωική οργάνωση.

29. Η γάτα του γείτονα επιτέθηκε στη δική μας γάτα όταν την πρωτοπήραμε, όμως δε θα μπορούσε να φανταστεί κανείς ότι θα γίνονταν οι καλύτεροι φίλοι μέσα σε δύο μήνες.

30. Ο ταχυδρόμος της γειτονιάς παρατηρήσαμε πως κάνει πολύ γρήγορα τη δουλειά του, και αναρωτιόμασταν μήπως προσπερνούσε κάποια σπίτια για να τελειώσει πιο γρήγορα.

31. Η Χριστίνα αγχωνόταν και έτρεχε πάνω κάτω μέσα στο σπίτι βιαστικά πριν το ραντεβού της, αλλά δεν μπορούσε να θυμηθεί πού είχε βάλει την τσάντα της και πού το πορτοφόλι της.

32. Η κοπέλα κοίταζε μελαγχολικά το κινητό περιμένοντας μήπως την πάρει τηλέφωνο ο αγαπημένος της, ωστόσο μετά από κάποια ώρα συνειδητοποίησε ότι περίμενε άδικα γιατί δεν είχε σήμα.

33. Οι καθοδηγημένες εξορμήσεις με μοτοσυκλέτες είναι μια νέα μορφή τουρισμού που προσελκύει αρκετό κόσμο, αλλά δεν μπορώ να φανταστώ ότι κάποιος από τους γνωστούς μας θα δελεαζόταν από αυτό.

34. Η κυβέρνηση κάνει κάθε προσπάθεια να περιορίσει τις εγκληματικές δραστηριότητες στη χώρα μας, όμως μελέτες δείχνουν πως αντιθέτως έχουν πολλαπλασιαστεί κατά την τελευταία δεκαετία.

35. Ο ζωγράφος της παρέας πήγαινε κάθε πρωί βόλτα στο πάρκο για έμπνευση, αλλά σήμερα εμπνεύστηκε από τον γάτο που καθόταν έξω από το παράθυρό του.

36. Η αστυνομία ανέκρινε την επιτροπή σχετικά με το ξέπλυμα μαύρου χρήματος, αλλά κανένας τους δεν ομολόγησε τίποτα και έτσι έπρεπε να πάει η υπόθεση στο δικαστήριο.

37. Ο αδερφός της Πολυξένης είχε ενθουσιαστεί πολύ με τους χειμερινούς Ολυμπιακούς αγώνες στην τηλεόραση, όμως οι υπόλοιποι δε γνωρίζαμε κανέναν από τους αθλητές και έτσι βαρεθήκαμε γρήγορα.

38. Η πωλήτρια μας έλεγε για τις τιμές των προϊόντων ομορφιάς και τις προσφορές, όμως ακόμα δεν κατάλαβα γιατί η κρέμα χεριών πωλείται πολύ πιο ακριβά από την κρέμα σώματος.

39. Ο καθηγητής μας έλεγε ότι δημιουργήθηκε ο πρώτος ψηφιακός χάρτης για τον αναπτυσσόμενο εγκέφαλο, και προκλήθηκε μεγάλος ενθουσιασμός στην επιστημονική κοινότητα ως συνέπεια αυτού του μεγάλου κατορθώματος.

40. Οι γονείς του κοριτσιού μάλωσαν έναν από τους συμμαθητές της όταν την έφεραν στο σχολείο, αλλά δεν ακούσαμε για ποιο λόγο τον μάλωσαν γιατί βρισκόμαστα αρκετά μακριά.

41. Η Πηνελόπη εξοργίστηκε όταν της είπαν οι συμφοιτητές της ότι έδωσαν τις σημειώσεις της κάπου, αλλά δεν μπόρεσε να εξακριβώσει πού τις έδωσαν παρόλο που προσπάθησε πολύ.

42. Ο υπαστυνόμος στη συνέντευξη τύπου είπε στους δημοσιογράφους ότι είχαν κάποιους υπόπτους, αλλά δεν αποκάλυψε ποιοι από αυτούς έπρεπε να διερευνηθούν περαιτέρω για αδιάσειστα ενοχοποιητικά στοιχεία.

43. Οι θεατές περίμεναν πώς και πώς να αποκαλυφθεί ο ένοχος στην τελευταία σκηνή της σειράς, αλλά τελικά δεν αποκαλύφθηκε από ποιον δολοφονήθηκε η αδερφή της πρωταγωνίστριας προς αγανάκτηση όλων.

44. Οι αρχαιολόγοι είχαν κάνει μεγάλες προόδους στις ανασκαφές της αρχαίας πόλης και νέα ιστορικά στοιχεία βγήκαν στο φως, ωστόσο παραδέχτηκαν ότι μπορεί ποτέ να μη μάθουμε πώς και πότε ακριβώς καταστράφηκε αυτή.

45. Οι εργαζόμενοι του νοσοκομείου είχαν κάνει ήδη πολλές διαμαρτυρίες και απεργίες, ωστόσο κανείς δεν ήξερε ακόμα πότε θα πληρώνονταν ή αν θα έμενε καν ανοιχτό το νοσοκομείο.

46. Τα λαμπραντόρ λέγεται ότι είναι η πιο υπάκουη ράτσα σκύλου, ωστόσο εμείς προτιμήσαμε να πάρουμε λυκόσκυλο διότι θέλαμε περισσότερο έναν φύλακα για το σπίτι.

47. Η γιαγιά του Μίλτου ήθελε χρόνια να υιοθετήσει μια γάτα από τη γειτονιά και τελικά πρόπερσι το έκανε, αλλά κάθε καλοκαίρι το μετανιώνει διότι μαδάει τόσο πολύ που θέλει συνέχεια καθάρισμα το σπίτι.

48. Η κόρη του Στέφανου από μικρό παιδί διάβαζε κάθε βιβλίο που έπιανε στα χέρια της, και έτσι δε θα μπορούσα να σου πω ποιο βιβλίο διαβάζει αυτή την περίοδο με απόλυτη σιγουριά.

49. Ο μικρός αδερφός της Λουίζας φοβάται πολύ τα φίδια, και οι γονείς του τον πήγανε σε ένα ζωολογικό κήπο να τα δει από κοντά ώστε να ξεπεράσει το φόβο του.

50. Κάθε παιδί λένε ότι ονειρεύεται να γίνει αστροναύτης, αστυνόμος ή πυροσβέστης όταν μεγαλώσει, όμως αναρωτιέμαι πόσα παιδιά ονειρεύονται να γίνουν μάγειρες ή βιβλιοθηκάριοι.

51. Η κηπουρός του ξενοδοχείου χαιρόταν πολύ να φροντίζει τα τριαντάφυλλα στην μπροστινή είσοδο, αλλά δεν ήταν ξεκάθαρο εάν είχε και δικό της προσωπικό κήπο στο σπίτι.

52. Ο μάγειρας του εστιατορίου έφτιαξε μία πεντανόστιμη πίτα για κυρίως πιάτο, και συζητούσαμε το πόσο ζηλεύαμε την οικογένειά του που έτρωγε τόσο ωραίο φαγητό καθημερινά.

53. Λένε ότι εάν βάλεις αλάτι στο νερό θα αργήσει περισσότερο να έρθει σε βράση, αλλά η ξαδέρφη μου ακόμα δεν μπορεί να καταλάβει γιατί συμβαίνει αυτό.

54. Ο γαμπρός της Τίνας κοίταζε αραιά και πού την ώρα για να μην αργήσει για τη συνάντηση, αλλά δεν κατάλαβε ότι το ρολόι του είχε σταματήσει πριν αρκετή ώρα και τελικά καθυστέρησε.

55. Η παιδίατρος είχε συμβουλέψει τη μαμά του Βασίλη να σταματήσει ο μικρός τα πατατάκια, αλλά κανείς δεν κατάλαβε ότι εκείνος τα αγόραζε κρυφά από το περίπτερο και τα έτρωγε στο σχολείο.

56. Η νύφη του Ανέστη λάτρευε το τρέξιμο και πήγαινε έξω στις έξι κάθε πρωί, αλλά κανείς δεν είχε αντιληφθεί ακόμα πού ακριβώς πήγαινε για να τρέξει.

57. Ο νταής της γειτονιάς έτρεξε κλαίγοντας στο σπίτι του χθες το βράδυ, αλλά ντρεπόταν να πει σε κανέναν ότι χτυπήθηκε από το μικρότερο κορίτσι της γειτονιάς.

58. Ο ανηψιός της Ραφαέλας είπε σε όλους ότι τραυματίστηκε σε μια σκληρή μάχη με έναν άγριο εχθρό, αλλά η αλήθεια είναι ότι γλίστρησε στο μπάνιο και έπεσε πάνω στο νεροχύτη.

59. Οι γριούλες της γειτονιάς μαζεύονταν κάθε Κυριακή για να παίξουν χαρτιά, αλλά δεν ήταν ξεκάθαρο ποια από όλες εξαπατώνταν περισσότερο από τις άλλες.

60. Το κτίριο στη γωνία λίγο πιο κάτω από το σπίτι μας είναι από τα παλαιότερα στην περιοχή, αλλά είναι μυστήριο από ποιον χτίστηκε ακριβώς και για ποιο λόγο.

61. Στο γιορτινό τραπέζι συζητούσαμε για τα διάφορα φαγητά που φτιάχνονται στον κόσμο, και αναρωτιόμασταν ποιο φαγητό μαγειρεύεται στη Βόρεια Ευρώπη για το Πάσχα.

62. Η μητέρα του Τοτού πληρώθηκε την προηγούμενη Παρασκευή ύστερα από αρκετό καιρό, και έτσι του αγόρασε ένα νέο παιχνίδι και στον εαυτό της έκανε δώρο ένα μανικιούρ.

63. Η καρέτα καρέτα ήταν είδος χελώνας υπό εξαφάνιση, αλλά επειδή προστατεύεται από το κράτος εδώ και πολλά χρόνια έχουν αυξηθεί αρκετά οι αριθμοί της.

64. Η γιαγιά του Αλκιβιάδη έφτιαξε τζατζίκι το σαββατοκύριακο για να φάνε τα εγγόνια, αλλά ήξερε ότι ακόμα κι αν δεν το φάνε διατηρείται εύκολα στο ψυγείο.

65. Η μικρή Νικολέτα έμαθε ότι η συμπεριφορά στο δρόμο ορίζεται από τον κώδικα οδικής κυκλοφορίας, αλλά δεν μπορούσε να καταλάβει πώς μπορεί να ορίζεται η οδήγηση του μπαμπά της ακριβώς.

66. Καθώς πηγαίναμε εκδρομή στον Ψηλορείτη πήραμε λάθος στροφή και χαθήκαμε, αλλά ευτυχώς βάλαμε το αυτόματο σύστημα πλοήγησης και καθοδηγηθήκαμε πάλι στο σωστό δρόμο.

67. Ο μικρότερος αδερφός της Άλκηστης είναι εξαιρετικά φιλομαθής αν και δεν έχει πάει δημοτικό ακόμα, και σήμερα τον βρήκαμε να διαβάζει πώς και πότε δημιουργήθηκε το διαδίκτυο πριν μερικά χρόνια.

68. Η κόρη του Σάκη είναι τριών χρονών και γεμάτη ιστορίες, και σήμερα μας έλεγε ότι κυνηγήθηκε από έναν μονόκερο στον κήπο μέχρι να φτάσει στο δωμάτιό της.

69. Ο ξάδερφος της Ελισάβετ αντιμετωπίζει πολλούς αντιπάλους στο γήπεδο μπάσκετ, αλλά δε θυμάμαι ποιανού αντιπάλου του εαυτού της σε κάθε αθλητικό αγώνα υπερτερεί η αδερφή του εδώ και χρόνια.

70. Ο αδερφός της Μαριάμ χάρηκε κάποια εγκώμια γραμμένα από την οικογένεια για τη γιορτή του, αλλά δεν ήξερε κανείς πόσων εγκωμίων για τον εαυτό της είχε αντίληψη η μαμά του για την δική της επερχόμενη γιορτή.

71. Η θεία του Τάκη ανακάλυψε κάποιες πτυχές του εαυτού της στο ταξίδι αναψυχής, αλλά δεν είμαστε σίγουροι ποιας πτυχής του εαυτού του στο ταξίδι απέκτησε αντίληψη ο αδερφός της πολύ γρήγορα.

72. Η νονά του Αρίσταρχου άγγιξε κάποια οικογενειακά θέματα στη συζήτηση, και ήταν πολύ φανερό ποιανού θέματος για τον εαυτό του στη συζήτηση αντιτάθηκε ο μπαμπάς του σθεναρά.

73. Ο μπαμπάς της Μαγδαληνής συζητούσε κάποιους ανταγωνιστές υπαλλήλων μέσα στην εταιρεία, και ήταν προφανές ποιανών ανταγωνιστών του εαυτού του υπερείχε στην εταιρεία ο συνάδελφός του σε κάθε συνεδρίαση.

74. Η μαμά του Ανδρέα μιλούσε για τους προγόνους της οικογένειας στο τραπέζι, αλλά δεν μπορούσαμε να καταλάβουμε ποιανού προγόνου του εαυτού της έμοιαζε η φίλη της που έκανε χρόνια παρέα μαζί της.

75. Ο ξάδερφος της Ελεονόρας συζητούσε για πολλούς καλούς φίλους του από το πανεπιστήμιο, αλλά ήταν ξεκάθαρο ποιανού φίλου του εαυτού του από τη σχολή έμοιαζε ο αδερφός του περισσότερο σε χαρακτήρα.

76. Η θεία του Λεωνίδα αγνοούσε πολλά ποιήματα του άνδρα της από τα νιάτα του, αλλά ξέραμε ποιανού ποιήματος για τον εαυτό της από παλιά είχε σίγουρα γνώση η μητέρα του και το είχε φυλάξει καλά.

77. Στις ειδήσεις το απόγευμα είπαν πως το χρηματιστήριο κινείται ανοδικά, αλλά στο σχολείο ακόμα δε μας εξήγησαν πώς επηρεάζεται ο απλός λαός από αυτό.

78. Το νεαρό ζευγάρι είχε επενδύσει σε μία μεγάλη έκταση γης για σπίτι, αλλά ήταν αμφίβολο εάν θα χτιζόταν το σπίτι μέσα στα επόμενα δύο χρόνια τουλάχιστον.

79. Τα παιδιά έψαχναν γεμάτα ενθουσιασμό τη γατούλα τους που κρυβόταν κάπου σήμερα το πρωί, αλλά δεν μπόρεσαν να καταλάβουν πού κρυβόταν παρόλο που έψαχναν για ώρες.

80. Ο πυρηνικός φυσικός έχει κάνει μέχρι τώρα αρκετές ανακοινώσεις σε σημαντικά επιστημονικά συνέδρια, και όλοι περιμένουν πότε θα αναγνωριστεί η δουλειά του με κάποιο βραβείο του κλάδου του.

81. Ο ξάδερφος της Μαρκετώς ήθελε να μάθει πάρα πολύ Γιαπωνέζικα, αλλά ήξερε ότι δεν θα μάθαινε εάν δεν πήγαινε κάπου που να μιλιέται αυτή η γλώσσα γύρω του.

82. Η μικρή κοπελίτσα πήγε στην Κίνα και αγόρασε μια σειρά εξωτικών τσαγιών, αλλά επειδή ήταν οι οδηγίες στα Κινέζικα δεν είχε ιδέα πώς προετοιμαζόταν κανένα από αυτά δυστυχώς.

83. Όλα τα παιδιά που έδιναν πανελλήνιες είχαν διαβάσει όσο το δυνατόν περισσότερο μπορούσαν, αλλά τα παιδιά από το δεύτερο ενιαίο λύκειο ήταν σίγουρα πως είχαν διαβάσει περισσότερο από όλους.

84. Η μητέρα του Ορέστη έφτιαξε κοτόπουλο με ρεβύθια στο φούρνο για μεσημεριανό, αλλά εκείνος δεν αισθανόταν καλά και έτσι δεν έφαγε καθόλου προς μεγάλη απογοήτευση της μητέρας του.

85. Ο παππούς της εξάχρονης ενθουσιάστηκε πολύ στη σχολική γιορτή, αλλά κανείς δε μπορούσε να φανταστεί πως θα χειροκροτούσε τόσο την εγγονή του που υποδυόταν έναν ασήμαντο ρόλο.

86. Η παρέα του γιού της μαζεύεται τακτικά στο σπίτι για να παίξουν ηλεκτρονικά παιχνίδια, και είναι κοινό μυστικό πως κανείς τους δεν αρέσει ιδιαίτερα στη μικρότερη αδελφή του.

87. Η καλύτερη φίλη της νηπιαγωγού έχει τη φήμη νευρωτικής μαμάς, αλλά κανείς δεν περίμενε πως θα απαιτούσε να ζητήσει η διευθύντρια συγγνώμη για τη συμπεριφορά της συμμαθήτριάς της.

88. Η συζήτηση για την προέλευση του σύμπαντος είναι παλιά όσο κι ο άνθρωπος, και δεν αναμένεται να απαντηθεί με πειστικό τρόπο από την επιστημονική κοινότητα στο προσεχές μέλλον.

89. Ο νόμος λέει ότι πρέπει όλοι να βάζουν μικροτσίπ στα κατοικίδια ζώα τους, ώστε σε περίπτωση που χαθούν να μπορούν εύκολα να βρεθούν οι ιδιοκτήτες από οποιονδήποτε κτηνίατρο.

90. Ο θείος του περιπτερά παρακολούθησε στο δημαρχείο μια μέθοδο βελτίωσης της μνήμης για ηλικιωμένους, αλλά ξέχασε ότι έπρεπε να επαναλαμβάνεται νοερά την ίδια μέρα και τα αποτελέσματα ήταν πενιχρά.

91. Στις πανελλήνιες φέτος στην έκθεση έπεσε το θέμα της φιλίας και της κοινωνικής δικτύωσης, και ο απουσιολόγος της τάξης χτυπιόταν επειδή ήταν το μόνο που δεν είχε κάνει επανάληψη την προηγούμενη ημέρα.

92. Ο μικρότερος αδερφός του κοριτσιού έσπασε την αγαπημένη της ροζ κούπα, ωστόσο εκείνη δεν έκλαψε επειδή ήξερε ότι η μαμά της θα της αγόραζε μία ακόμη καλύτερη και μεγαλύτερη μπλε κούπα.

93. Η πρωτεύουσα της γειτονικής χώρας είναι φημισμένη για τα κοσμοπολίτικα εστιατόριά της, και όλοι ενθουσιάστηκαν όταν ανακοινώθηκε η απόφαση να πραγματοποιηθεί εκεί η ετήσια εκδρομή σε τρείς μήνες.

94. Υποτίθεται πως το να αποφεύγεις τα πολλά γλυκά βοηθά στο να χάσεις βάρος, αλλά δεν βοηθά το γεγονός ότι ο άντρας της υπέρβαρης ξαδέλφης της δουλεύει σε ζαχαροπλαστείο που δεν τον πληρώνει τακτικά.

95. Ο καθηγητής της Λήδας είχε ακούσε κάποιες φήμες για το προσωπικό στο γραφείο, αλλά δεν ήξερε κανείς πόσες φήμες για τον εαυτό της από το γραφείο είχε ακούσει η νέα διευθύντρια από τότε που έφτασε.

96. Τα τριαντάφυλλα άρεσαν πολύ στη μαμά του κοριτσιού, όμως ο μπαμπάς του κοριτσιού ήθελε να της κάνει δώρο ένα πανέμορφο μπουκέτο κόκκινες τουλίπες για αλλαγή.

97. Η θετική σκέψη είναι το πρώτο βήμα για την αλλαγή προς το καλύτερο, και η μικρή Αννούλα το ήξερε πολύ καλά αυτό αν και βαριόταν να το εφαρμόσει.

98. Η αρχιτέκτονας έφτιαξε τα νέα σχέδια για το σπίτι που της είχαν παραγγείλει οι ιδιοκτήτες, όμως είχε άγχος μήπως δεν τους αρέσουν παρόλο που είχε περάσει όλη τη βδομάδα δουλεύοντας από το πρωί μέχρι το βράδυ.

99. Ο Τοτός άκουσε πως ο ταχυδρόμος χτυπάει πάντα δυο φορές, αλλά ήταν σίγουρος πως ο δικός τους ταχυδρόμος δεν είχε χτυπήσει ούτε μια φορά στο σπίτι τους εδώ και χρόνια.

100. Η ανηψιά του Κώστα είχε ως χόμπι να μαζεύει πολύ όμορφες ξυλομπογιές, ωστόσο το παράξενο ήταν ότι ποτέ δεν τις χρησιμοποιούσε αλλά τις φύλαγε σε ένα κουτί για να τις θαυμάζει.

C.2 Simplified English SPR Study

C.2.1 Residual RT Mean and SEM Results per Region per Condition

Table C.3: English SPR RT Means and SEM per Region per Condition

but

	Condition	RT Resid. Mean	SEM	Log10 Resid. Mean	SEM
1	Wh-NP, P-less Verb	49.685	16.528	0.105	0.028
2	Wh-PP, P-less Verb	39.557	14.560	0.090	0.028
3	Wh-NP, PP-Verb	24.869	18.446	0.057	0.030
4	Wh-PP, PP-Verb	22.313	19.018	0.048	0.034

I

	Condition	RT Resid. Mean	SEM	Log10 Resid. Mean	SEM
1	Wh-NP, P-less Verb	-32.193	7.107	-0.082	0.018
2	Wh-PP, P-less Verb	-25.252	8.922	-0.067	0.020
3	Wh-NP, PP-Verb	-35.555	6.852	-0.097	0.019
4	Wh-PP, PP-Verb	-39.695	7.089	-0.099	0.020

do

	Condition	RT Resid. Mean	SEM	Log10 Resid. Mean	SEM
1	Wh-NP, P-less Verb	-14.569	6.652	-0.042	0.015
2	Wh-PP, P-less Verb	-16.041	6.874	-0.041	0.018
3	Wh-NP, PP-Verb	-18.881	8.527	-0.072	0.016
4	Wh-PP, PP-Verb	-21.581	8.327	-0.059	0.020

not

	Condition	RT Resid. Mean	SEM	Log10 Resid. Mean	SEM
1	Wh-NP, P-less Verb	-10.597	5.604	-0.025	0.014
2	Wh-PP, P-less Verb	-9.179	5.630	-0.027	0.016
3	Wh-NP, PP-Verb	-10.329	7.115	-0.034	0.018
4	Wh-PP, PP-Verb	-8.407	5.451	-0.032	0.015

know

	Condition	RT Resid. Mean	SEM	Log10 Resid. Mean	SEM
1	Wh-NP, P-less Verb	-7.200	6.577	-0.020	0.017
2	Wh-PP, P-less Verb	-7.646	6.550	-0.021	0.017
3	Wh-NP, PP-Verb	-18.362	6.002	-0.050	0.017
4	Wh-PP, PP-Verb	-21.728	6.167	-0.057	0.019

for.sure

	Condition	RT Resid. Mean	SEM	Log10 Resid. Mean	SEM
1	Wh-NP, P-less Verb	0.103	10.738	0.007	0.028
2	Wh-PP, P-less Verb	6.741	11.870	0.024	0.028
3	Wh-NP, PP-Verb	-3.953	10.448	-0.003	0.028
4	Wh-PP, PP-Verb	-15.162	11.923	-0.033	0.031

to

	Condition	RT Resid. Mean	SEM	Log10 Resid. Mean	SEM
1	Wh-NP, P-less Verb	N/A	N/A	N/A	N/A
2	Wh-PP, P-less Verb	27.270	7.561	0.052	0.016
3	Wh-NP, PP-Verb	N/A	N/A	N/A	N/A
4	Wh-PP, PP-Verb	1.142	5.919	-0.007	0.017

which.rumours

	Condition	RT Resid. Mean	SEM	Log10 Resid. Mean	SEM
1	Wh-NP, P-less Verb	-14.174	13.655	-0.042	0.028
2	Wh-PP, P-less Verb	-78.470	10.990	-0.167	0.028
3	Wh-NP, PP-Verb	-27.874	8.771	-0.063	0.023
4	Wh-PP, PP-Verb	-90.279	12.669	-0.213	0.033

Annabelle

	Condition	RT Resid. Mean	SEM	Log10 Resid. Mean	SEM
1	Wh-NP, P-less Verb	-27.866	16.188	-0.086	0.024
2	Wh-PP, P-less Verb	12.619	23.078	-0.007	0.034
3	Wh-NP, PP-Verb	-37.565	19.938	-0.109	0.031
4	Wh-PP, PP-Verb	-1.559	17.981	-0.025	0.027

responded

	Condition	RT Resid. Mean	SEM	Log10 Resid. Mean	SEM
1	Wh-NP, P-less Verb	57.212	16.625	0.137	0.034
2	Wh-PP, P-less Verb	72.034	16.612	0.173	0.033
3	Wh-NP, PP-Verb	45.814	14.771	0.113	0.032
4	Wh-PP, PP-Verb	60.540	13.838	0.161	0.032

to

	Condition	RT Resid. Mean	SEM	Log10 Resid. Mean	SEM
1	Wh-NP, P-less Verb	22.529	8.121	0.062	0.014
2	Wh-PP, P-less Verb	N/A	N/A	N/A	N/A
3	Wh-NP, PP-Verb	36.449	8.382	0.096	0.017
4	Wh-PP, PP-Verb	N/A	N/A	N/A	N/A

and

	Condition	RT Resid. Mean	SEM	Log10 Resid. Mean	SEM
1	Wh-NP, P-less Verb	7.569	5.780	0.026	0.015
2	Wh-PP, P-less Verb	49.557	10.589	0.118	0.017
3	Wh-NP, PP-Verb	14.477	7.757	0.030	0.015
4	Wh-PP, PP-Verb	43.633	12.499	0.097	0.018

I

	Condition	RT Resid. Mean	SEM	Log10 Resid. Mean	SEM
1	Wh-NP, P-less Verb	7.302	6.755	0.012	0.015
2	Wh-PP, P-less Verb	30.524	7.530	0.075	0.018
3	Wh-NP, PP-Verb	0.330	5.852	-0.005	0.015
4	Wh-PP, PP-Verb	24.833	6.919	0.062	0.016

would.like.to.know

	Condition	RT Resid. Mean	SEM	Log10 Resid. Mean	SEM
1	Wh-NP, P-less Verb	20.495	31.824	0.072	0.074
2	Wh-PP, P-less Verb	63.709	27.859	0.214	0.073
3	Wh-NP, PP-Verb	-16.323	31.563	-0.010	0.073
4	Wh-PP, PP-Verb	51.638	40.141	0.136	0.085

C.2.2 Raw RT Mean and SEM Results per Region per Condition

but

	Condition	Raw RT Mean	SEM	Log10 RT Mean	SEM
1	Wh-NP, P-less Verb	464.869	25.209	6.030	0.046
2	Wh-PP, P-less Verb	454.741	22.470	6.015	0.044
3	Wh-NP, PP-Verb	440.053	25.323	5.982	0.045
4	Wh-PP, PP-Verb	437.497	24.686	5.973	0.048

I

	Condition	Raw RT Mean	SEM	Log10 RT Mean	SEM
1	Wh-NP, P-less Verb	330.694	15.003	5.735	0.041
2	Wh-PP, P-less Verb	337.634	14.621	5.750	0.040
3	Wh-NP, PP-Verb	327.331	16.285	5.720	0.043
4	Wh-PP, PP-Verb	323.198	14.682	5.718	0.042

do

	Condition	Raw RT Mean	SEM	Log10 RT Mean	SEM
1	Wh-NP, P-less Verb	315.522	15.502	5.684	0.041
2	Wh-PP, P-less Verb	314.050	15.312	5.685	0.041
3	Wh-NP, PP-Verb	311.209	16.347	5.655	0.042
4	Wh-PP, PP-Verb	308.509	13.781	5.667	0.039

not

	Condition	Raw RT Mean	SEM	Log10 RT Mean	SEM
1	Wh-NP, P-less Verb	309.206	12.070	5.680	0.037
2	Wh-PP, P-less Verb	310.625	13.720	5.678	0.040
3	Wh-NP, PP-Verb	309.475	13.342	5.671	0.038
4	Wh-PP, PP-Verb	311.397	13.999	5.673	0.040

know

	Condition	Raw RT Mean	SEM	Log10 RT Mean	SEM
1	Wh-NP, P-less Verb	319.906	14.075	5.697	0.041
2	Wh-PP, P-less Verb	319.459	15.439	5.696	0.043
3	Wh-NP, PP-Verb	308.744	14.026	5.667	0.042
4	Wh-PP, PP-Verb	305.378	13.763	5.660	0.041

for.sure

	Condition	Raw RT Mean	SEM	Log10 RT Mean	SEM
1	Wh-NP, P-less Verb	657.009	28.535	11.456	0.080
2	Wh-PP, P-less Verb	663.647	31.322	11.472	0.082
3	Wh-NP, PP-Verb	652.953	28.840	11.446	0.081
4	Wh-PP, PP-Verb	641.744	26.508	11.416	0.078

to

	Condition	Raw RT Mean	SEM	Log10 RT Mean	SEM
1	Wh-NP, P-less Verb	N/A	N/A	N/A	N/A
2	Wh-PP, P-less Verb	352.888	17.149	5.781	0.042
3	Wh-NP, PP-Verb	N/A	N/A	N/A	N/A
4	Wh-PP, PP-Verb	326.759	15.248	5.722	0.041

which.rumours

	Condition	Raw RT Mean	SEM	Log10 RT Mean	SEM
1	Wh-NP, P-less Verb	673.138	30.663	11.487	0.080
2	Wh-PP, P-less Verb	672.884	30.598	11.502	0.082
3	Wh-NP, PP-Verb	659.438	28.429	11.465	0.080
4	Wh-PP, PP-Verb	661.075	28.178	11.456	0.079

Annabelle

	Condition	Raw RT Mean	SEM	Log10 RT Mean	SEM
1	Wh-NP, P-less Verb	390.356	26.392	5.833	0.051
2	Wh-PP, P-less Verb	408.118	33.794	5.863	0.062
3	Wh-NP, PP-Verb	380.636	30.307	5.810	0.058
4	Wh-PP, PP-Verb	393.950	28.549	5.845	0.054

responded

	Condition	Raw RT Mean	SEM	Log10 RT Mean	SEM
1	Wh-NP, P-less Verb	596.136	26.967	8.526	0.085
2	Wh-PP, P-less Verb	583.484	27.942	8.480	0.086
3	Wh-NP, PP-Verb	583.469	27.153	8.477	0.094
4	Wh-PP, PP-Verb	574.681	25.801	8.489	0.093

to

	Condition	Raw RT Mean	SEM	Log10 RT Mean	SEM
1	Wh-NP, P-less Verb	364.369	18.209	5.827	0.041
2	Wh-PP, P-less Verb	N/A	N/A	N/A	N/A
3	Wh-NP, PP-Verb	378.307	17.291	5.862	0.040
4	Wh-PP, PP-Verb	N/A	N/A	N/A	N/A

and

	Condition	Raw RT Mean	SEM	Log10 RT Mean	SEM
1	Wh-NP, P-less Verb	339.882	14.397	5.771	0.040
2	Wh-PP, P-less Verb	381.844	21.195	5.863	0.042
3	Wh-NP, PP-Verb	346.784	18.681	5.774	0.043
4	Wh-PP, PP-Verb	375.941	22.475	5.842	0.045

I

	Condition	Raw RT Mean	SEM	Log10 RT Mean	SEM
1	Wh-NP, P-less Verb	328.934	12.735	5.738	0.035
2	Wh-PP, P-less Verb	352.156	15.039	5.801	0.038
3	Wh-NP, PP-Verb	321.963	13.128	5.721	0.038
4	Wh-PP, PP-Verb	346.466	15.052	5.788	0.039

would.like.to.know

	Condition	Raw RT Mean	SEM	Log10 RT Mean	SEM
1	Wh-NP, P-less Verb	1668.581	62.648	26.853	0.158
2	Wh-PP, P-less Verb	1716.475	66.716	27.036	0.185
3	Wh-NP, PP-Verb	1633.750	58.583	26.775	0.180
4	Wh-PP, PP-Verb	1702.491	65.173	26.923	0.192

C.2.3 English SPR Stimuli

Experimental Items

Conditions are only described for the first item.

1. i. **Condition 1: Wh-NP; P-less Verb**
   Oliver heard some rumours at the pub, but I do not know for sure which rumours Annabelle responded to and I would like to know.

   ii. **Condition 2: Wh-PP; P-less Verb**
   Oliver heard some rumours at the pub, but I do not know for sure to which rumours Annabelle responded and I would like to know.

   iii. **Condition 3: Wh-NP; PP-Verb**
   Oliver listened to some rumours at the pub, but I do not know for sure which rumours Annabelle responded to and I would like to know.

   iv. **Condition 4: Wh-PP; PP-Verb**
   Oliver listened to some rumours at the pub, but I do not know for sure to which rumours Annabelle responded and I would like to know.

2. i. Oliver heard some rumours at the pub, but I do not know for sure which rumours Annabelle responded to and I would like to know.

   ii. Oliver heard some rumours at the pub, but I do not know for sure to which rumours Annabelle responded and I would like to know.

iii. Oliver listened to some rumours at the pub, but I do not know for sure which rumours Annabelle responded to and I would like to know.

iv. Oliver listened to some rumours at the pub, but I do not know for sure to which rumours Annabelle responded and I would like to know.

3. i. Julia saw some pictures on the internet, but we did not know for sure which pictures Christopher reacted to and we were keen to find out.

   ii. Julia saw some pictures on the internet, but we did not know for sure to which pictures Christopher reacted and we were keen to find out.

   iii. Julia objected to some pictures on the internet, but we did not know for sure which pictures Christopher reacted to and we were keen to find out.

   iv. Julia objected to some pictures on the internet, but we did not know for sure to which pictures Christopher reacted and we were keen to find out.

4. i. Adam sold some drawings at the gallery, but I did not see very clearly which drawings Stephanie was pleased with and I did not care either.

   ii. Adam sold some drawings at the gallery, but I did not see very clearly with which drawings Stephanie was pleased and I did not care either.

   iii. Adam was impressed with some drawings at the gallery, but I did not see very clearly which drawings Stephanie was pleased with and I did not care either.

   iv. Adam was impressed with some drawings at the gallery, but I did not see very clearly with which drawings Stephanie was pleased and I did not care either.

5. i. Chloe drew some sketches during art class, but we could not tell for sure which sketches Harrison was happy with and we wanted to find out.

   ii. Chloe drew some sketches during art class, but we could not tell for sure with which sketches Harrison was happy and we wanted to find out.

   iii. Chloe was pleased with some sketches during art class, but we could not tell for sure which sketches Harrison was happy with and we wanted to find out.

   iv. Chloe was pleased with some sketches during art class, but we could not tell for sure with which sketches Harrison was happy and we wanted to find out.

6. i. Andrew admired some presents at the dinner, but I could not say for sure which presents Catherine was annoyed with and I was eager to know.

   ii. Andrew admired some presents at the dinner, but I could not say for sure with which presents Catherine was annoyed and I was eager to know.

   iii. Andrew was pleased with some presents at the dinner, but I could not say for sure which presents Catherine was annoyed with and I was eager to know.

iv. Andrew was pleased with some presents at the dinner, but I could not say for sure with which presents Catherine was annoyed and I was eager to know.

7. i. Rose heard some songs on the radio, but I did not realise straight away which songs Sebastian was impressed with and I was surprised to find out.

   ii. Rose heard some songs on the radio, but I did not realise straight away with which songs Sebastian was impressed and I was surprised to find out.

   iii. Rose was annoyed with some songs on the radio, but I did not realise straight away which songs Sebastian was impressed with and I was surprised to find out.

   iv. Rose was annoyed with some songs on the radio, but I did not realise straight away with which songs Sebastian was impressed and I was surprised to find out.

8. i. Samuel heard some stories at the nursery, but I did not hear very clearly which stories Genevieve reacted to and I am curious to know.

   ii. Samuel heard some stories at the nursery, but I did not hear very clearly to which stories Genevieve reacted and I am curious to know.

   iii. Samuel listened to some stories at the nursery, but I did not hear very clearly which stories Genevieve reacted to and I am curious to know.

   iv. Samuel listened to some stories at the nursery, but I did not hear very clearly to which stories Genevieve reacted and I am curious to know.

9. i. Alice took some photos in the restaurant, but I could not see too clearly which photos Nathaniel laughed at and I was not that interested either.

   ii. Alice took some photos in the restaurant, but I could not see too clearly at which photos Nathaniel laughed and I was not that interested either.

   iii. Alice stared at some photos in the restaurant, but I could not see too clearly which photos Nathaniel laughed at and I was not that interested either.

   iv. Alice stared at some photos in the restaurant, but I could not see too clearly at which photos Nathaniel laughed and I was not that interested either.

10. i. Nick drew some sketches in the studio, but we did not hear for sure which sketches Magdalena was impressed with and we were interested to know.

    ii. Nick drew some sketches in the studio, but we did not hear for sure with which sketches Magdalena was impressed and we were interested to know.

    iii. Nick was happy with some sketches in the studio, but we did not hear for sure which sketches Magdalena was impressed with and we were interested to know.

    iv. Nick was happy with some sketches in the studio, but we did not hear for sure with which sketches Magdalena was impressed and we were interested to know.

11. i. Lucy noticed some presents at the party, but I could not tell for sure which presents Demetrius was pleased with and I was keen to find out.
    ii. Lucy noticed some presents at the party, but I could not tell for sure with which presents Demetrius was pleased and I was keen to find out.
    iii. Lucy was disappointed with some presents at the party, but I could not tell for sure which presents Demetrius was pleased with and I was keen to find out.
    iv. Lucy was disappointed with some presents at the party, but I could not tell for sure with which presents Demetrius was pleased and I was keen to find out.
12. i. Stephen liked some photos on his camera, but we did not realise straight away which photos Alexandra was appalled at and we were dying to find out.
    ii. Stephen liked some photos on his camera, but we did not realise straight away at which photos Alexandra was appalled and we were dying to find out.
    iii. Stephen laughed at some photos on his camera, but we did not realise straight away which photos Alexandra was appalled at and we were dying to find out.
    iv. Stephen laughed at some photos on his camera, but we did not realise straight away at which photos Alexandra was appalled and we were dying to find out.
13. i. Amanda heard some rumours in the office, but we could not tell with certainty which rumours Theodore objected to and we were keen to find out.
    ii. Amanda heard some rumours in the office, but we could not tell with certainty to which rumours Theodore objected and we were keen to find out.
    iii. Amanda contributed to some rumours in the office, but we could not tell with certainty which rumours Theodore objected to and we were keen to find out.
    iv. Amanda contributed to some rumours in the office, but we could not tell with certainty to which rumours Theodore objected and we were keen to find out.
14. i. Hugo heard some stories on the news, but I did not understand with certainty which stories Hermione laughed at and I was not that interested either.
    ii. Hugo heard some stories on the news, but I did not understand with certainty at which stories Hermione laughed and I was not that interested either.
    iii. Hugo marvelled at some stories on the news, but I did not understand with certainty which stories Hermione laughed at and I was not that interested either.
    iv. Hugo marvelled at some stories on the news, but I did not understand with certainty at which stories Hermione laughed and I was not that interested either.
15. i. Amber admired some drawings at the competition, but I could not remember with clarity which drawings Anderson voted for and I am curious to know.

ii. Amber admired some drawings at the competition, but I could not remember with clarity for which drawings Anderson voted and I am curious to know.

iii. Amber was rooting for some drawings at the competition, but I could not remember with clarity which drawings Anderson voted for and I am curious to know.

iv. Amber was rooting for some drawings at the competition, but I could not remember with clarity for which drawings Anderson voted and I am curious to know.

16. i. Lucas found some pictures in the attic, but we did not realise straight away which pictures Charlotte laughed at and we were keen to find out.

ii. Lucas found some pictures in the attic, but we did not realise straight away at which pictures Charlotte laughed and we were keen to find out.

iii. Lucas marvelled at some pictures in the attic, but we did not realise straight away which pictures Charlotte laughed at and we were keen to find out.

iv. Lucas marvelled at some pictures in the attic, but we did not realise straight away at which pictures Charlotte laughed and we were keen to find out.

17. i. Jessica played some games at the party, but we could not recall with certainty which games Alexander voted for and we were surprised to find out.

ii. Jessica played some games at the party, but we could not recall with certainty for which games Alexander voted and we were surprised to find out.

iii. Jessica asked for some games at the party, but we could not recall with certainty which games Alexander voted for and we were surprised to find out.

iv. Jessica asked for some games at the party, but we could not recall with certainty for which games Alexander voted and we were surprised to find out.

18. i. James took some courses at the college, but I did not realise straight away which courses Anastasia objected to and I was surprised to find out.

ii. James took some courses at the college, but I did not realise straight away to which courses Anastasia objected and I was surprised to find out.

iii. James contributed to some courses at the college, but I did not realise straight away which courses Anastasia objected to and I was surprised to find out.

iv. James contributed to some courses at the college, but I did not realise straight away to which courses Anastasia objected and I was surprised to find out.

19. i. Phoebe repeated some questions from the seminar, but I could not tell for sure which questions Montgomery was annoyed with and I would like to know.

ii. Phoebe repeated some questions from the seminar, but I could not tell for sure with which questions Montgomery was annoyed and I would like to know.

iii. Phoebe agreed with some questions from the seminar, but I could not tell for sure which questions Montgomery was annoyed with and I would like to know.

iv. Phoebe agreed with some questions from the seminar, but I could not tell for sure with which questions Montgomery was annoyed and I would like to know.

20. i. Oscar watched some films on the television, but I did not understand for sure which films Bernadette was appalled at and I am waiting to find out.

ii. Oscar watched some films on the television, but I did not understand for sure at which films Bernadette was appalled and I am waiting to find out.

iii. Oscar laughed at some films on the television, but I did not understand for sure which films Bernadette was appalled at and I am waiting to find out.

iv. Oscar laughed at some films on the television, but I did not understand for sure at which films Bernadette was appalled and I am waiting to find out.

21. i. Holly watched some programmes on the television, but I could not tell with certainty which programmes Christian was waiting for and I was curious to know.

ii. Holly watched some programmes on the television, but I could not tell with certainty for which programmes Christian was waiting and I was curious to know.

iii. Holly paid for some programmes on the television, but I could not tell with certainty which programmes Christian was waiting for and I was curious to know.

iv. Holly paid for some programmes on the television, but I could not tell with certainty for which programmes Christian was waiting and I was curious to know.

22. i. Robert liked some books at the auction, but we could not recall very well which books Gabrielle paid for and we were astonished to find out.

ii. Robert liked some books at the auction, but we could not recall very well for which books Gabrielle paid and we were astonished to find out.

iii. Robert bidded for some books at the auction, but we could not recall very well which books Gabrielle paid for and we were astonished to find out.

iv. Robert bidded for some books at the auction, but we could not recall very well for which books Gabrielle paid and we were astonished to find out.

23. i. Laura played some games in the park, but I did not hear very clearly which games Bartholomew marvelled at and I am not that interested either.

ii. Laura played some games in the park, but I did not hear very clearly at which games Bartholomew marvelled and I am not that interested either.

iii. Laura laughed at some games in the park, but I did not hear very clearly which games Bartholomew marvelled at and I am not that interested either.

iv. Laura laughed at some games in the park, but I did not hear very clearly at which games Bartholomew marvelled and I am not that interested either.

24. i. William saw some boats near the dock, but we could not recall very well which boats Josephine marvelled at and we were not that interested either.
 ii. William saw some boats near the dock, but we could not recall very well at which boats Josephine marvelled and we were not that interested either.
 iii. William stared at some boats near the dock, but we could not recall very well which boats Josephine marvelled at and we were not that interested either.
 iv. William stared at some boats near the dock, but we could not recall very well at which boats Josephine marvelled and we were not that interested either.

25. i. Katie supported some films at the festival, but I could not say with certainty which films Maximillian voted for and I would be interested to know.
 ii. Katie supported some films at the festival, but I could not say with certainty for which films Maximillian voted and I would be interested to know.
 iii. Katie waited for some films at the festival, but I could not say with certainty which films Maximillian voted for and I would be interested to know.
 iv. Katie waited for some films at the festival, but I could not say with certainty for which films Maximillian voted and I would be interested to know.

26. i. Lewis recognised some problems on the blackboard, but I do not remember very well which problems Elizabeth laughed at and I was not very interested either.
 ii. Lewis recognised some problems on the blackboard, but I do not remember very well at which problems Elizabeth laughed and I was not very interested either.
 iii. Lewis stared at some problems on the blackboard, but I do not remember very well which problems Elizabeth laughed at and I was not very interested either.
 iv. Lewis stared at some problems on the blackboard, but I do not remember very well at which problems Elizabeth laughed and I was not very interested either.

27. i. Anna noticed some problems in the study, but I do not recall with certainty which problems Benjamin was prepared for and I was surprised to find out.
 ii. Anna noticed some problems in the study, but I do not recall with certainty for which problems Benjamin was prepared and I was surprised to find out.
 iii. Anna apologised for some problems in the study, but I do not recall with certainty which problems Benjamin was prepared for and I was surprised to find out.
 iv. Anna apologised for some problems in the study, but I do not recall with certainty for which problems Benjamin was prepared and I was surprised to find out.

28. i. Edward played some songs on the piano, but I do not recall with certainty which songs Valentina reacted to and I would like to know.
    ii. Edward played some songs on the piano, but I do not recall with certainty to which songs Valentina reacted and I would like to know.
    iii. Edward listened to some songs on the piano, but I do not recall with certainty which songs Valentina reacted to and I would like to know.
    iv. Edward listened to some songs on the piano, but I do not recall with certainty to which songs Valentina reacted and I would like to know.

29. i. Emily spotted some boats on the lake, but I cannot say with certainty which boats Johnathan was looking for and I would be interested to know.
    ii. Emily spotted some boats on the lake, but I cannot say with certainty for which boats Johnathan was looking and I would be interested to know.
    iii. Emily waited for some boats on the lake, but I cannot say with certainty which boats Johnathan was looking for and I would be interested to know.
    iv. Emily waited for some boats on the lake, but I cannot say with certainty for which boats Johnathan was looking and I would be interested to know.

30. i. John attended some courses at the college, but I do not remember very clearly which courses Jacqueline was happy with and I do not care either.
    ii. John attended some courses at the college, but I do not remember very clearly with which courses Jacqueline was happy and I do not care either.
    iii. John was impressed with some courses at the college, but I do not remember very clearly which courses Jacqueline was happy with and I do not care either.
    iv. John was impressed with some courses at the college, but I do not remember very clearly with which courses Jacqueline was happy and I do not care either.

31. i. Victoria watched some programmes on the television, but I could not tell for sure which programmes Frederic objected to and I was eager to find out.
    ii. Victoria watched some programmes on the television, but I could not tell for sure to which programmes Frederic objected and I was eager to find out.
    iii. Victoria listened to some programmes on the television, but I could not tell for sure which programmes Frederic objected to and I was eager to find out.
    iv. Victoria listened to some programmes on the television, but I could not tell for sure to which programmes Frederic objected and I was eager to find out.

32. i. Jack answered some questions in the exam, but I did not realise straight away which questions Evangeline was prepared for and I was astonished to find out.

ii. Jack answered some questions in the exam, but I did not realise straight away for which questions Evangeline was prepared and I was astonished to find out.

iii. Jack hoped for some questions in the exam, but I did not realise straight away which questions Evangeline was prepared for and I was astonished to find out.

iv. Jack hoped for some questions in the exam, but I did not realise straight away for which questions Evangeline was prepared and I was astonished to find out.

33. i. Natalie found some books at the bookshop, but I could not tell with certainty which books Dominic paid for and I was keen to know.

ii. Natalie found some books at the bookshop, but I could not tell with certainty for which books Dominic paid and I was keen to know.

iii. Natalie asked for some books at the bookshop, but I could not tell with certainty which books Dominic paid for and I was keen to know.

iv. Natalie asked for some books at the bookshop, but I could not tell with certainty for which books Dominic paid and I was keen to know.

Fillers

1. Roger told us some stories at the dinner, but I do not remember too well which stories and I would like to find out somehow.

2. Leah reprimanded the children for some stories, which they had told in the playground because they had offended people.

3. Mary wanted to take some pictures from the mantelpiece, but I did not see very clearly which pictures and I'm not all that interested either.

4. Joshua heard some rumours at the dinner party, which his mother hosted and couldn't wait to tell us about them.

5. Oscar really liked some pictures from the family album, but I cannot remember right now which pictures and I would like to find out somehow.

6. Roger sold some drawings at the gallery event which he had worked on for months and was very pleased with the selling price.

7. Brandon drew some sketches at the museum, but I did not see for sure how many sketches and I wasn't all that interested either.

8. Audrey admired some drawings at the fair, which was being held in her home town and decided to buy a few of them.

9. Ronan picked up some photographs which he had stored in the old house and was excited to show them to his family.

10. Simon drew some sketches for the comic, but I could not see very clearly how many sketches and I would be interested to find out.

11. Ruby took some photographs in the room, which she had carefully planned the lighting in and was delighted with the results.

12. Christian told people about some rumours from the cafeteria, but I did not hear with certainty which rumours and I'm dying to find out.

13. Helen showed everyone some photographs on the screen, but I did not see very clearly which photographs and I am feeling left out.

14. Penny told everyone about some rumours from the office, but we did not understand too well how many rumours and we did not really care either.

15. Lara drew some sketches in her portfolio, which she had been practising for the comic book and was still not happy with the results.

16. Neil pointed out some drawings on the wall, but we could not see very clearly which drawings and no one cared enough to ask either.

17. Poppy told some stories in the playground, but I could not hear too clearly which stories and I would like to find out.

18. Louisa chose some sketches for the fashion article, but no one knew with certainty which sketches and everyone was dying to know.

19. Caleb spoke about some photographs at the art show, but I could not figure out for sure which photographs and I was puzzled.

20. Jasmine went on about some courses at the college, but I do not remember right now how many courses and to be honest I do not really care.

21. Elsie brought many presents to the birthday party, but I could not count with certainty how many presents and I am curious to know.

22. Timothy bartered for some books in the shop, but we did not understand right away how many books and we were shocked when we found out.

23. Hannah brought some games to summer camp, but I did not know right away how many games and was very pleased when I found out.

24. Grace discussed some problems during maths class, but I could not hear very clearly which problems and I would definitely like to know.

25. Walter wanted to watch some films after the dinner, but no one could figure out for sure which films and they did not care either.

26. Ruth needed help with some questions in the exam, but I could not figure out for sure which questions and so I could not help her.

27. Ian really loved some songs on the radio, but we did not know for sure which songs and needed to find out for his surprise party.

28. Nicole liked some boats on the lake, but I did not see very clearly which boats and would like to find out to photograph them for her.

29. Samantha wanted to watch some programmes on the television, but I did not see with certainty which programmes and I would like to know so I can join her.

30. Dylan taught some courses at the university, but no one could remember with certainty which courses and they were all too embarrassed to admit it.

31. Noah brought some presents for the housewarming, but I could not remember too well which presents and needed to ask him.

32. Sylvia brought some books on the camping trip, but I could not have imagined at the time how many books and was shocked to find out.

33. Aaron played some games in the playground, but we did not know for sure how many games and were not that interested either.

34. Andrew presented some problems in the meeting, but I did not hear with certainty how many problems and now I wish I had.

35. Edith made some films during art class, but no one knew with certainty how many films and we were very impressed when we found out.

36. Basil asked some questions during the seminar, but I did not hear very clearly how many questions and was amused when I found out.

37. Elsa recorded some songs in the studio, but I did not know for sure how many songs and was very surprised when I found out.

38. Ernest owned some boats in the boat yard, but I never knew with certainty how many boats and I plan to ask him soon.

39. Willow followed lots of programmes on the radio, which she had inherited from her grandfather, because she liked to learn by listening.

40. Callum liked some courses at the local college, which had excellent lecturers, and decided to enrol thanks to them.

41. Louise bought some presents on the wedding list, which she found out about at the last minute, and was happy she still managed to get cheap ones.

42. Owen wanted some books from the bookshop, which was having a closing down sale and was delighted when his mother bought them for him.

43. Rose wanted to play some games at summer camp, which she attended every year, and was trying to find someone to play them with her.

44. Gary was worried about some problems in his architectural design, which he had been working on for months, and wondered how to fix them.

45. Anne watched some films at the cinema, which she and her parents always visit, and liked them so much that she decided to see them again.

46. Neville answered some questions on the exam, which was particularly hard and was convinced that he had failed miserably.

47. Lucas sang some songs at the party, which he had been practising for and everyone was very impressed by his voice.

48. Elizabeth brought some drawings to the class, but I did not see right away how many drawings and was impressed when I found out.

49. Mary was very happy with some brushes, which she bought at the art store, and was eager to show her classmates.

50. Magda was bullied by George at school, but she did not realise until later that he actually liked her.

51. Henry was looking forward to going to the doctor's, since his hay fever was getting particularly bad and he needed medicine.

52. Amanda edited some manuscripts in the early afternoon, since she was planning on going to the cinema later with her boyfriend.

53. Tony decided to bring the teacher some gifts, but was disappointed when his marks still did not improve and his friends found it hilarious.

54. Mary waited for her friend at school, but he was absent from several classes that day and she was worried that something had happened to him.

55. Larry was placed in charge of the department, but not everyone felt he was the most capable at handling all of the administrative work.

56. Harry was admired by several girls in his year, but he was completely oblivious to this and his friends envied him greatly.

57. Nicola was awarded a medal of honour for her courage , but she felt that she still had a lot more to give the service.

58. Charlie successfully negotiated the technical contract, but it was his people skills and not his presentation which won the day.

59. Celia was finally made a partner at the law firm, but she cared more about the better hours than the increase in salary.

60. Hazel had flirted with Nicholas at the firm party, but he still did not realise that she liked him more than just as a friend.

61. Luke was investigating some electoral fraud allegations, but was dismayed to learn that the journalists had already found out about them.

62. Ruth watched some young athletes at the tryouts, but the expert talent scout took too long to judge them and so she left.

63. Aaron was watched by his family in his new stage performance, but no one had expected him to really be that good.

64. Ivy tidied up the house in the afternoon, but it was her amazing cooking that really impressed her family and friends.

C.2.4 English SC Stimuli

Experimental Items

Conditions are only described for the first item.

1. i. **Condition 1: Wh-NP, P-less Verb**
   Oliver heard some rumours at the pub, but I did not know for sure ____

   ii. **Condition 2: Wh-PP, P-less Verb**
   Oliver heard some rumours at the pub, but I did not know for sure to ____

   iii. **Condition 3: Wh-NP, PP-Verb**
   Oliver listened to some rumours at the pub, but I did not know for sure ____

iv. **Condition 4: Wh-PP, PP-Verb** Oliver listened to some rumours at the pub, but I did not know for sure to ____

2. i. Julia saw some pictures on the internet, but we did not know for sure ____
   ii. Julia saw some pictures on the internet, but we did not know for sure to ____
   iii. Julia objected to some pictures on the internet, but we did not know for sure ____
   iv. Julia objected to some pictures on the internet, but we did not know for sure to ____

3. i. Adam sold some drawings at the gallery, but I did not see very clearly ____
   ii. Adam sold some drawings at the gallery, but I did not see very clearly with ____
   iii. Adam was impressed with some drawings at the gallery, but I did not see very clearly ____
   iv. Adam was impressed with some drawings at the gallery, but I did not see very clearly with ____

4. i. Chloe drew some sketches during art class, but we could not tell for sure ____
   ii. Chloe drew some sketches during art class, but we could not tell for sure with ____
   iii. Chloe was pleased with some sketches during art class, but we could not tell for sure ____
   iv. Chloe was pleased with some sketches during art class, but we could not tell for sure with ____

5. i. Andrew admired some presents at the dinner, but I could not say for sure ____
   ii. Andrew admired some presents at the dinner, but I could not say for sure with ____
   iii. Andrew was pleased with some presents at the dinner, but I could not say for sure ____
   iv. Andrew was pleased with some presents at the dinner, but I could not say for sure with ____

6. i. Rose heard some songs on the radio, but I did not realise straight away ____
   ii. Rose heard some songs on the radio, but I did not realise straight away with ____
   iii. Rose was annoyed with some songs on the radio, but I did not realise straight away ____
   iv. Rose was annoyed with some songs on the radio, but I did not realise straight away with ____

7. i. Samuel heard some stories at the nursery, but I did not hear very clearly ____
   ii. Samuel heard some stories at the nursery, but I did not hear very clearly to ____
   iii. Samuel listened to some stories at the nursery, but I did not hear very clearly ____
   iv. Samuel listened to some stories at the nursery, but I did not hear very clearly to ____
8. i. Alice took some photos in the restaurant, but I could not see too clearly ____
   ii. Alice took some photos in the restaurant, but I could not see too clearly at ____
   iii. Alice stared at some photos in the restaurant, but I could not see too clearly ____
   iv. Alice stared at some photos in the restaurant, but I could not see too clearly at ____
9. i. Nick drew some sketches in the studio, but we did not hear for sure ____
   ii. Nick drew some sketches in the studio, but we did not hear for sure with ____
   iii. Nick was happy with some sketches in the studio, but we did not hear for sure ____
   iv. Nick was happy with some sketches in the studio, but we did not hear for sure with ____
10. i. Lucy noticed some presents at the party, but I could not tell for sure ____
    ii. Lucy noticed some presents at the party, but I could not tell for sure with ____
    iii. Lucy was disappointed with some presents at the party, but I could not tell for sure ____
    iv. Lucy was disappointed with some presents at the party, but I could not tell for sure with ____
11. i. Stephen liked some photos on his camera, but we did not realise straight away ____
    ii. Stephen liked some photos on his camera, but we did not realise straight away at ____
    iii. Stephen laughed at some photos on his camera, but we did not realise straight away ____
    iv. Stephen laughed at some photos on his camera, but we did not realise straight away at ____
12. i. Amanda heard some rumours in the office, but we could not tell with certainty ____

ii. Amanda heard some rumours in the office, but we could not tell with certainty to ____

iii. Amanda contributed to some rumours in the office, but we could not tell with certainty ____

iv. Amanda contributed to some rumours in the office, but we could not tell with certainty to ____

13. i. Hugo heard some stories on the news, but I did not understand with certainty ____

ii. Hugo heard some stories on the news, but I did not understand with certainty at ____

iii. Hugo marvelled at some stories on the news, but I did not understand with certainty ____

iv. Hugo marvelled at some stories on the news, but I did not understand with certainty at ____

14. i. Amber admired some drawings at the competition, but I could not remember with clarity ____

ii. Amber admired some drawings at the competition, but I could not remember with clarity for ____

iii. Amber was rooting for some drawings at the competition, but I could not remember with clarity ____

iv. Amber was rooting for some drawings at the competition, but I could not remember with clarity for ____

15. i. Lucas found some pictures in the attic, but we did not realise straight away ____

ii. Lucas found some pictures in the attic, but we did not realise straight away at ____

iii. Lucas marvelled at some pictures in the attic, but we did not realise straight away ____

iv. Lucas marvelled at some pictures in the attic, but we did not realise straight away at ____

16. i. Jessica played some games at the party, but we could not recall with certainty ____

ii. Jessica played some games at the party, but we could not recall with certainty for ____

iii. Jessica asked for some games at the party, but we could not recall with certainty ____

iv. Jessica asked for some games at the party, but we could not recall with certainty for ____

17. i. James took some courses at the college, but I did not realise straight away ____

    ii. James took some courses at the college, but I did not realise straight away to ____

    iii. James contributed to some courses at the college, but I did not realise straight away ____

    iv. James contributed to some courses at the college, but I did not realise straight away to ____

18. i. Phoebe repeated some questions from the seminar, but I could not tell for sure ____

    ii. Phoebe repeated some questions from the seminar, but I could not tell for sure with ____

    iii. Phoebe agreed with some questions from the seminar, but I could not tell for sure ____

    iv. Phoebe agreed with some questions from the seminar, but I could not tell for sure with ____

19. i. Oscar watched some films on the television, but I did not understand for sure ____

    ii. Oscar watched some films on the television, but I did not understand for sure at ____

    iii. Oscar laughed at some films on the television, but I did not understand for sure ____

    iv. Oscar laughed at some films on the television, but I did not understand for sure at ____

20. i. Holly watched some programmes on the television, but I could not tell with certainty ____

    ii. Holly watched some programmes on the television, but I could not tell with certainty for ____

    iii. Holly paid for some programmes on the television, but I could not tell with certainty ____

    iv. Holly paid for some programmes on the television, but I could not tell with certainty for ____

21. i. Robert liked some books at the auction, but we could not recall very well ____
    ii. Robert liked some books at the auction, but we could not recall very well for ____
    iii. Robert bidded for some books at the auction, but we could not recall very well ____
    iv. Robert bidded for some books at the auction, but we could not recall very well for ____
22. i. Laura played some games in the park, but I did not hear very clearly ____
    ii. Laura played some games in the park, but I did not hear very clearly at ____
    iii. Laura laughed at some games in the park, but I did not hear very clearly ____
    iv. Laura laughed at some games in the park, but I did not hear very clearly at ____
23. i. William saw some boats near the dock, but we could not recall very well ____
    ii. William saw some boats near the dock, but we could not recall very well at ____
    iii. William stared at some boats near the dock, but we could not recall very well ____
    iv. William stared at some boats near the dock, but we could not recall very well at ____
24. i. Katie supported some films at the festival, but I could not say with certainty ____
    ii. Katie supported some films at the festival, but I could not say with certainty for ____
    iii. Katie waited for some films at the festival, but I could not say with certainty ____
    iv. Katie waited for some films at the festival, but I could not say with certainty for ____
25. i. Lewis recognised some problems on the blackboard, but I do not remember very well ____
    ii. Lewis recognised some problems on the blackboard, but I do not remember very well at ____
    iii. Lewis stared at some problems on the blackboard, but I do not remember very well ____
    iv. Lewis stared at some problems on the blackboard, but I do not remember very well at ____
26. i. Anna noticed some problems in the study, but I do not recall with certainty ____

ii. Anna noticed some problems in the study, but I do not recall with certainty for ____

iii. Anna apologised for some problems in the study, but I do not recall with certainty ____

iv. Anna apologised for some problems in the study, but I do not recall with certainty for ____

27. i. Edward played some songs on the piano, but I do not recall with certainty ____

ii. Edward played some songs on the piano, but I do not recall with certainty to ____

iii. Edward listened to some songs on the piano, but I do not recall with certainty ____

iv. Edward listened to some songs on the piano, but I do not recall with certainty to ____

28. i. Emily spotted some boats on the lake, but I cannot say with certainty ____

ii. Emily spotted some boats on the lake, but I cannot say with certainty for ____

iii. Emily waited for some boats on the lake, but I cannot say with certainty ____

iv. Emily waited for some boats on the lake, but I cannot say with certainty for ____

29. i. John attended some courses at the college, but I do not remember very clearly ____

ii. John attended some courses at the college, but I do not remember very clearly with ____

iii. John was impressed with some courses at the college, but I do not remember very clearly ____

iv. John was impressed with some courses at the college, but I do not remember very clearly with ____

30. i. Victoria watched some programmes on the television, but I could not tell for sure ____

ii. Victoria watched some programmes on the television, but I could not tell for sure to ____

iii. Victoria listened to some programmes on the television, but I could not tell for sure ____

iv. Victoria listened to some programmes on the television, but I could not tell for sure to ____

31. i. Jack answered some questions in the exam, but I did not realise straight away ____
    ii. Jack answered some questions in the exam, but I did not realise straight away for ____
    iii. Jack hoped for some questions in the exam, but I did not realise straight away ____
    iv. Jack hoped for some questions in the exam, but I did not realise straight away for ____
32. i. Natalie found some books at the bookshop, but I could not tell with certainty ____
    ii. Natalie found some books at the bookshop, but I could not tell with certainty for ____
    iii. Natalie asked for some books at the bookshop, but I could not tell with certainty ____
    iv. Natalie asked for some books at the bookshop, but I could not tell with certainty for ____

Fillers

1. Roger told us some stories at the dinner that were all about heroes saving ____
2. Leah reprimanded the children for some stories, which they had told in the playground because ____
3. Mary wanted to take some pictures from the mantelpiece that were photographs of ____
4. Joshua heard some rumours at the dinner party, which his mother hosted and ____
5. Oscar really liked some pictures from the family album, particularly the photographs that reminded ____
6. Roger sold some drawings at the gallery event which he had worked on for months and ____
7. Brandon drew some sketches at the museum and his classmates were amused ____
8. Audrey admired some drawings at the fair, which was being held in her home town and ____
9. Ronan picked up some photographs which he had stored in the old house and ____

10. Simon drew some sketches for the comic book and he would like to send ____
11. Ruby took some photographs in the room, which she had carefully planned the lighting in and ____
12. Christian told people about some rumours from the cafeteria,, but no one trusted ____
13. Helen showed everyone some photographs on the screen and all of them appreciated ____
14. Penny told everyone about some rumours from the office, but no one believed ____
15. Lara drew some sketches in her portfolio, which she had been practising for the comic book and ____
16. Neil pointed out some drawings on the wall and he tried to explain ____
17. Poppy told some stories in the playground that were all about ____
18. Lousia chose some sketches for the fashion article and everyone thought ____
19. Caleb spoke about some photographs at the art show and he was amazed ____
20. Jasmine went on about some courses at the college that she struggled to ____
21. Elsie brought many presents to the birthday party, they were all packed ____
22. Timothy bartered for some books in the shop and was very proud of ____
23. Hannah brought some games to summer camp that her friends had ____
24. Grace discussed some problems during maths class that were so ____
25. Walter wanted to watch some films after the dinner, so he asked his friends ____
26. Ruth needed help with some questions in the exam, so she emailed her professor for ____
27. Ian really loved some songs on the radio, so he sang along while he was ____
28. Nicole liked some boats on the lake that were colourful and ____
29. Samantha wanted to watch some programmes on the television so she sat down on the sofa with ____
30. Dylan taught some courses at the university and he was planning to ____
31. Noah brought some presents for the housewarming and he spent ____

32. Sylvia brought some books on the camping trip and while she managed to finish one book ____
33. Aaron played some games in the playground and he had a lot of ____
34. Andrew presented some problems in the meeting, and while everyone tried hard to find a solution ____
35. Edith made some films during art class and her teacher was impressed ____
36. Basil asked some questions during the seminar, as he would like to ____
37. Elsa recorded some songs in the studio and she sent them ____
38. Ernest owned some boats in the boat yard and he was planning to ____
39. Willow followed lots of programmes on the radio, which she had inherited from her grandfather, because ____
40. Callum liked some courses at the local college, which had excellent lecturers, and ____
41. Louise bought some presents on the wedding list, which she found out about at the last minute, and ____
42. Owen wanted some books from the bookshop, which was having a closing down sale and ____
43. Rose wanted to play some games at summer camp, which she attended every year, and ____
44. Gary was worried about some problems in his architectural design, which he had been working on for months, and ____
45. Anne watched some films at the cinema, which she and her parents always visit, and ____
46. Neville answered some questions on the exam, which was particularly hard and ____
47. Lucas sang some songs at the party, which he had been practising for and ____
48. Elizabeth brought some drawings to the class and they were going to be presented ____
49. Mary was very happy with some brushes, which she bought at the art store, and ____
50. Magda was bullied by George at school,, but she did not realise until later that ____

51. Henry was looking forward to going to the doctor's, since his hay fever was getting particularly bad and ____
52. Amanda edited some manuscripts in the early afternoon, since ____
53. Tony decided to bring the teacher some gifts, but was disappointed ____
54. Mary waited for her friend at school, but he was absent from several classes that day and ____
55. Larry was placed in charge of the department, but ____
56. Harry was admired by several girls in his year, but he was completely oblivious to this and ____
57. Nicola was awarded a medal of honour for her courage , but she felt that ____
58. Charlie successfully negotiated the technical contract, but it was his people skills and ____
59. Celia was finally made a partner at the law firm, but ____
60. Hazel had flirted with Nicholas at the firm party, but he still did not realise that ____
61. Luke was investigating some electoral fraud allegations, but was dismayed to learn that ____
62. Ruth watched some young athletes at the tryouts, but ____
63. Aaron was watched by his family in his new stage performance, but ____
64. Ivy tidied up the house in the afternoon, but it was her amazing cooking that ____

References

Abels, K. (2003). *Successive cyclicity, anti-locality, and adposition stranding* (Unpublished doctoral dissertation). University of Connecticut.

Abels, K. (2011). Don't fix that island! it ain't broke. In *Islands in contemporary linguistic theory conference, vitoria-gasteiz.*

Abels, K. (2016). *On the interaction of p-stranding and sluicing in bulgarian.* Ms. UCL.

Abels, K. (2017). Movement and islands. *The Oxford handbook of ellipsis*.

Abels, K., & Dayal, V. (2017). On the syntax of multiple sluicing. In *Proceedings of nels 47.*

Abels, K., & Thoms, G. (2014). Resumption and island repair under sluicing. In *Lagb annual meeting.*

Albert, C. (1993). Sluicing and weak islands. *Ms., University of California, Santa Cruz*.

Alexiadou, A., & Anagnostopoulou, E. (2000). Asymmetries in the distribution of clitics. *Clitic phenomena in European languages*, *30*, 47.

Almeida, D. A. d. A., & Yoshida, M. (2007). A problem for the preposition stranding generalization. *Linguistic Inquiry*, *38*(2), 349–362.

Almor, A. (1999). Noun-phrase anaphora and focus: The informational load hypothesis. *Psychological review*, *106*(4), 748.

Almor, A., Kempler, D., MacDonald, M. C., Andersen, E. S., & Tyler, L. K. (1999). Why do alzheimer patients have difficulty with pronouns? working memory, semantics, and reference in comprehension and production in alzheimer's disease. *Brain and language*, *67*(3), 202–227.

AlShaalan, Y., & Abels, K. (2019). Resumption as a sluicing source in saudi arabic: evidence from apparent p-stranding. *Glossa*.

Altmann, G., & Steedman, M. (1988). Interaction with context during human sentence processing. *Cognition*, *30*(3), 191–238.

Altmann, G. T., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, *73*(3), 247–264.

Alwin, D. F., & Krosnick, J. A. (1991). The reliability of survey attitude measurement the influence of question and respondent attributes. *Sociological Methods & Research*, *20*(1), 139–181.

Ambridge, B., Pine, J. M., Rowland, C. F., & Young, C. R. (2008). The effect of verb semantic class and verb frequency (entrenchment) on children's and adults' graded judgements of argument-structure overgeneralization errors. *Cognition*, *106*(1), 87–129.

Aoun, J., Hornstein, N., Lightfoot, D., & Weinberg, A. (1987). Two types of locality. *Linguistic inquiry*, *18*(4), 537–577.

Ariel, M. (1990). *Accessing np antecedents.* London: Routledge.

Ariel, M. (1994). Interpreting anaphoric expressions: A cognitive versus a pragmatic approach. *Journal of linguistics*, *30*(1), 3–42.

Ariel, M. (2001). Accessibility theory: An overview. *Text representation: Linguistic and psycholinguistic aspects*, *8*, 29–87.

Ariel, M. (2014). *Accessing noun-phrase antecedents (rle linguistics b: Grammar).* Routledge.

Arnon, I., & Snider, N. (2010). More than words: Frequency effects for multi-word phrases. *Journal of memory and language*, *62*(1), 67–82.

Arregui, A., Clifton, C., Frazier, L., & Moulton, K. (2006). Processing elided verb phrases with flawed antecedents: The recycling hypothesis. *Journal of memory and language*, *55*(2), 232–246.

Badecker, W., & Straub, K. (2002). The processing role of structural constraints on interpretation of pronouns and anaphors. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(4), 748.

Bader, M., & Häussler, J. (2010). Toward a model of grammaticality judgments. *Journal of Linguistics*, *46*(2), 273–330.

Bader, M., Meng, M., & Bayer, J. (2000). Case and reanalysis. *Journal of Psycholinguistic Research*, *29*(1), 37–52.

Baker, C. L., & Brame, M. K. (1972). 'global rules': a rejoinder. *Language*, 51–75.

Balota, D. A., Pollatsek, A., & Rayner, K. (1985). The interaction of contextual constraints and parafoveal visual information in reading. *Cognitive psychology*, *17*(3), 364–390.

Bard, E. G., Robertson, D., & Sorace, A. (1996). Magnitude estimation of linguistic acceptability. *Language*, 32–68.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of memory and language*, *68*(3), 255–278.

Barros, M., Elliott, P., & Thoms, G. (2014). There is no island repair. *Ms., Rutgers, University College London, and University of Edinburgh*.

Barros, M. V. (2013). Harmonic sluicing: Which remnant/correlate pairs work and why. In *Semantics and linguistic theory* (Vol. 23, pp. 295–315).

Bastiaanse, R., Rispens, J., Ruigendijk, E., Rabadán, O. J., & Thompson, C. K. (2002). Verbs: some properties and their consequences for agrammatic broca's aphasia. *Journal of Neurolinguistics*, *15*(3-5), 239–264.

Beck, S., & Johnson, K. (2004). Double objects again. *Linguistic Inquiry*, *35*(1), 97–123.

Bermel, N., & Knittl, L. (2012). Corpus frequency and acceptability judgments: A study of morphosyntactic variants in czech.

Birkett, N. J. (1986). Selecting the number of response categories for a likert-type scale. In *Proceedings of the american statistical association* (pp. 488–492).

Bock, K., & Miller, C. A. (1991). Broken agreement. *Cognitive psychology*, *23*(1), 45–93.

Boeckx, C. (2008). *Bare syntax*. OUP Oxford.

Bradlow, A. R., Kraus, N., & Hayes, E. (2003). Speaking clearly for children with learning disabilities: Sentence perception in noise. *Journal of Speech, Language, and Hearing Research*, *46*(1), 80–97.

Bregman, A. S. (1994). *Auditory scene analysis: The perceptual organization of sound*. MIT press.

Brysbaert, M., Mandera, P., & Keuleers, E. (2018). The word frequency effect in word processing: An updated review. *Current Directions in Psychological Science*, *27*(1), 45–50.

Cable, S., & Harris, J. A. (2011). On the grammatical status of pp-pied-piping in english: Results from sentence-rating experiments. *University of Massachusetts Occasional Papers in Linguistics*, *38*, 122.

Caha, P. (2011). Case in adpositional phrases. *Ms., CASTL, Tromsø*, 1–29.

Cai, Q., & Brysbaert, M. (2010). Subtlex-ch: Chinese word and character frequencies based on film subtitles. *PloS one*, *5*(6), e10729.

Caplan, D., & Waters, G. (2013). Memory mechanisms supporting syntactic comprehension. *Psychonomic bulletin & review*, *20*(2), 243–268.

Caplan, D., & Waters, G. S. (1999). Verbal working memory and sentence comprehension. *Behavioral and brain Sciences*, *22*(1), 77–94.

Caramazza, A., Grober, E., Garvey, C., & Yates, J. (1977). Comprehension of anaphoric pronouns. *Journal of verbal learning and verbal behavior*, *16*(5), 601–609.

Caramazza, A., & Zurif, E. B. (1976). Dissociation of algorithmic and heuristic processes in language comprehension: Evidence from aphasia. *Brain and language*, *3*(4), 572–582.

Chao, W. (1987). *On ellipsis, unpublished ph. d* (Unpublished doctoral dissertation). dissertation, University of Massachusetts, Amherst.

Chatsiou, A. (2006). On the status of resumptive pronouns in modern greek restrictive relative clauses. In *Proceedings of the lfg06 conference, university of konstanz*.

Chen, E. H., & Dixon, W. (1972). Estimates of parameters of a censored regression sample. *Journal of the American Statistical Association*, *67*(339), 664–671.

Chomsky, N. (1966). *Current issues in linguistic theory* (Vol. 38). Mouton.

Chomsky, N. (1972). *Some empirical issues in the theory of transformational grammar. goals of linguistic theory, ed. by stanley peters, 63-130.* Englewood Cliffs, NJ: Prentice-Hall.

Chomsky, N. (1981). *Lectures on government and binding. dordrecht: Foris.-. 1986a. knowledge of language: Its nature, origin, and use.* New York: Praeger.-. 1986b. Barriers. Cambridge, MA: MIT Press.

Chung, S., Ladusaw, W. A., & McCloskey, J. (1995). Sluicing and logical form. *Natural language semantics*, *3*(3), 239–282.

Cinque, G. (1990). *Types of ā-dependencies.* MIT press.

Collins, C. (2007). Home sweet home. *NYU WPL*, *1*, 1–34.

Criss, A. H., & McClelland, J. L. (2006). Differentiating the differentiation models: A comparison of the retrieving effectively from memory model (rem) and the subjective likelihood model (slim). *Journal of Memory and Language*, *55*(4), 447–460.

Crocker, M. W., & Brants, T. (2000). Wide-coverage probabilistic sentence processing. *Journal of Psycholinguistic Research*, *29*(6), 647–669.

Culicover, P. W., & Jackendoff, R. (2005). Simpler syntax.

Dąbrowska, E. (2010). Naive v. expert intuitions: An empirical study of acceptability judgments. *The linguistic review*, *27*(1), 1–23.

Dalrymple, M., Shieber, S. M., & Pereira, F. C. (1991). Ellipsis and higher-order unification. *Linguistics and philosophy*, *14*(4), 399–452.

Dayal, V., & Schwarzschild, R. (2010). Definite inner antecedents and wh-correlates in sluices. *Rutgers working papers in linguistics*, *3*, 92–114.

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature neuroscience*, *8*(8), 1117.

Demberg, V., & Keller, F. (2008). Data from eye-tracking corpora as evidence for theories of syntactic processing complexity. *Cognition*, *109*(2), 193–210.

Demberg, V., Keller, F., & Koller, A. (2013). Incremental, predictive parsing with psycholinguistically motivated tree-adjoining grammar. *Computational Linguistics*, *39*(4), 1025–1066.

De Ruiter, J.-P., Mitterer, H., & Enfield, N. J. (2006). Projecting the end of a speaker's turn: A cognitive cornerstone of conversation. *Language*, *82*(3), 515–535.

Dimitropoulou, M., Duñabeitia, J. A., Avilés, A., Corral, J., & Carreiras, M. (2010). Subtitle-based word frequencies as the best estimate of reading behavior: The case of greek. *Frontiers in psychology*, *1*, 218.

Divjak, D. (2008). On (in) frequency and (un) acceptability. *Corpus Linguistics, Computer Tools and Applications-state of the art*, 213–233.

Donaldson, T. S. (1968). Robustness of the f-test to errors of both kinds and the correlation between the numerator and denominator of the f-ratio. *Journal of the American Statistical Association*, *63*(322), 660–676.

Doya, K., Ishii, S., Pouget, A., & Rao, R. P. (2007). *Bayesian brain: Probabilistic approaches to neural coding.* MIT press.

Eberhard, K. M., Cutting, J. C., & Bock, K. (2005). Making syntax of sense: number agreement in sentence production. *Psychological Review*, *112*(3), 531.

Eguren, L. (2010). Contrastive focus and nominal ellipsis in spanish. *Lingua*, *120*(2), 435–457.

Ehrlich, K. (1980). Comprehension of pronouns. *The Quarterly Journal of Experimental Psychology*, *32*(2), 247–255.

Elman, J. L. (1990). Finding structure in time. *Cognitive science*, *14*(2), 179–211.

Embick, D., Hackl, M., Schaeffer, J., Kelepir, M., & Marantz, A. (2001). A magnetoencephalographic component whose latency reflects lexical frequency. *Cognitive Brain Research*, *10*(3), 345–348.

Erteschik-Shir, N. (1977). *On the nature of island constraints.* Indiana University Linguistics Club.

Featherston, S. (2008). Thermometer judgments as linguistic evidence. *Was ist linguistische Evidenz*, 69–89.

Featherston, S. (2009). A scale for measuring well-formedness: Why syntax needs boiling and freezing points. *The fruits of empirical linguistics*, *1*, 47–74.

Federmeier, K. D., Wlotko, E. W., De Ochoa-Dewald, E., & Kutas, M. (2007). Multiple effects of sentential constraint on word processing. *Brain research*, *1146*, 75–84.

Ferguson, C. J. (2009). An effect size primer: A guide for clinicians and researchers. *Professional Psychology: Research and Practice*, *40*(5), 532.

Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive psychology*, *47*(2), 164–203.

Ferreira, F., Christianson, K., & Hollingworth, A. (2001). Misinterpretations of garden-path sentences: Implications for models of sentence processing and reanalysis. *Journal of psycholinguistic research*, *30*(1), 3–20.

Ferreira, F., & Clifton Jr, C. (1986). The independence of syntactic processing. *Journal of memory and language*, *25*(3), 348–368.

Ferreira, F., & Patson, N. D. (2007). The 'good enough'approach to language comprehension. *Language and Linguistics Compass*, *1*(1-2), 71–83.

Fiengo, R., & May, R. (1994). *Indices and identity* (Vol. 24). MIT press.

Fodor, J., Bever, A., Garrett, T., et al. (1974). The psychology of language: An introduction to psycholinguistics and generative grammar.

Ford, M., Bresnan, J., & Kaplan, R. (n.d.). M.(1982). a competence–based theory of syntactic closure. *The mental representation of grammatical relations*, 727–796.

Forster, K. I., & Chambers, S. M. (1973). Lexical access and naming time. *Journal of Memory and Language*, *12*(6), 627.

Forster, K. I., & Hector, J. (2002). Cascaded versus noncascaded models of lexical and semantic processing: Theturple effect. *Memory & cognition*, *30*(7), 1106–1117.

Forster, K. I., & Shen, D. (1996). No enemies in the neighborhood: Absence of inhibitory neighborhood effects in lexical decision and semantic categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(3), 696.

Fox, D., & Lasnik, H. (2003). Successive-cyclic movement and island repair: The difference between sluicing and vp-ellipsis. *Linguistic inquiry*, *34*(1), 143–154.

Francis, W. N. (1986). Proximity concord in english. *Journal of English Linguistics*, *19*(2), 309–317.

Franck, J., Lassi, G., Frauenfelder, U. H., & Rizzi, L. (2006). Agreement and movement: A syntactic analysis of attraction. *Cognition*, *101*(1), 173–216.

Franck, J., Soare, G., Frauenfelder, U. H., & Rizzi, L. (2010). Object interference in subject–verb agreement: The role of intermediate traces of movement. *Journal of Memory and Language*, *62*(2), 166–182.

Frank, S. L. (2010). Uncertainty reduction as a measure of cognitive processing effort. In *Proceedings of the 2010 workshop on cognitive modeling and computational linguistics* (pp. 81–89).

Frank, S. L., & Bod, R. (2011). Insensitivity of the human sentence-processing system to hierarchical structure. *Psychological science*, *22*(6), 829–834.

Frank, S. L., Otten, L. J., Galli, G., & Vigliocco, G. (2013). Word surprisal predicts n400 amplitude during reading. In *Proceedings of the 51st annual meeting of the association for computational linguistics (volume 2: Short papers)* (Vol. 2, pp. 878–883).

Frank, S. L., Otten, L. J., Galli, G., & Vigliocco, G. (2015). The erp response to the amount of information conveyed by words in sentences. *Brain and language*, *140*, 1–11.

Frazier, L. (1978). *On comprehending sentences: Syntactic parsing strategies. university of connecticut ph. d* (Unpublished doctoral dissertation). dissertation, distributed by Indiana University Linguistics Club.

Frazier, L., & Fodor, J. D. (1978). The sausage machine: A new two-stage parsing model. *Cognition*, *6*(4), 291–325.

Freyd, M. (1923). The graphic rating scale. *Journal of Educational Psychology*, *14*(2), 83.

Friedmann, N., & Shapiro, L. P. (2003). Agrammatic comprehension of simple active sentences with moved constituentshebrew osv and ovs structures. *Journal of Speech, Language, and Hearing Research*, *46*(2), 288–297.

Fukaya, T. (2007). *Sluicing and stripping in japanese and some implications.* University of Southern California.

Fukaya, T. (2012). Island-sensitivity in japanese sluicing and some implications". *Sluicing:*

Cross-linguistic perspectives, *38*, 123.

Garnsey, S. M., Pearlmutter, N. J., Myers, E., & Lotocky, M. A. (1997). The contributions of verb bias and plausibility to the comprehension of temporarily ambiguous sentences. *Journal of memory and language*, *37*(1), 58–93.

Garnsey, S. M., Tanenhaus, M. K., & Chapman, R. M. (1989). Evoked potentials and the study of sentence comprehension. *Journal of psycholinguistic research*, *18*(1), 51–60.

Garrod, S., Freudenthal, D., & Boyle, E. (1994). The role of different types of anaphor in the on-line resolution of. *Journal of memory and language*, *33*, 39–68.

Garrod, S., & Pickering, M. J. (2015). The use of content and timing to predict turn transitions. *Frontiers in psychology*, *6*, 751.

Gibson, E. (1991). *A computational theory of human linguistic processing: Memory limitations and processing breakdown* (Unpublished doctoral dissertation). School of Computer Science, Carnegie Mellon University.

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, *68*(1), 1–76.

Gibson, E., Bergen, L., & Piantadosi, S. T. (2013). Rational integration of noisy evidence and prior semantic expectations in sentence interpretation. *Proceedings of the National Academy of Sciences*, 201216438.

Ginzburg, J., & Sag, I. (2000). *Interrogative investigations. csli.* Stanford.

Glass, G. V., Peckham, P. D., & Sanders, J. R. (1972). Consequences of failure to meet assumptions underlying the fixed effects analyses of variance and covariance. *Review of educational research*, *42*(3), 237–288.

Gordon, P. C., Grosz, B. J., & Gilliom, L. A. (1993). Pronouns, names, and the centering of attention in discourse. *Cognitive science*, *17*(3), 311–347.

Gordon, P. C., Hendrick, R., Ledoux, K., & Yang, C. L. (1999). Processing of reference and the structure of language: An analysis of complex noun phrases. *Language and Cognitive Processes*, *14*(4), 353–379.

Gorrell, P. (1989). Establishing the loci of serial and parallel effects in syntactic processing. *Journal of psycholinguistic research*, *18*(1), 61–73.

Gorrell, P. G. (1987). Studies of human syntactic processing: Ranked-parallel versus serial models.

Green, G. M. (1974). *Semantics and syntactic regularity.* Indiana Univ Pr.

Gribanova, V. (2013). Copular clauses, clefts, and putative sluicing in uzbek. *Language*, 830–882.

Griffiths, J., & Lipták, A. (2014). Contrast and island sensitivity in clausal ellipsis. *Syntax*, *17*(3), 189–234.

Gussenhoven, C. (1984). *On the grammar and semantics of sentence accents.* Dordrecht: Foris.

Hale, J. (2001). A probabilistic earley parser as a psycholinguistic model. In *Proceedings of the second meeting of the north american chapter of the association for computational linguistics on language technologies* (pp. 1–8).

Hale, J. (2003). The information conveyed by words in sentences. *Journal of Psycholinguistic Research*, *32*(2), 101–123.

Hale, J. T. (2011). What a rational parser would do. *Cognitive Science*, *35*(3), 399–443.

Hardt, D. (1993). Verb phrase ellipsis: Form, meaning, and processing.

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In *Psychology of learning and motivation* (Vol. 22, pp. 193–225). Elsevier.

Häussler, J., & Bader, M. (2009). Agreement checking and number attraction in sentence comprehension: Insights from german relative clauses. *Trav. Cercle Linguist. Prague*, *7*.

Hayes, M. H., & Patterson, D. G. (1921). Experimental development of the graphic rating method. *Psychol Bull*, *18*(1), 98–9.

Heck, F. (2008). *On pied-piping: Wh-movement and beyond* (Vol. 98). Walter de Gruyter.

Hofmeister, P. (2007). Memory retrieval effects on filler-gap processing. In *Proceedings of the 29th annual meeting of the cognitive science society* (pp. 1091–1096).

Hofmeister, P. (2011). Representational complexity and memory retrieval in language comprehension. *Language and cognitive processes*, *26*(3), 376–405.

Hofmeister, P., Casasanto, L. S., & Sag, I. A. (2014). Processing effects in linguistic judgment data:(super-) additivity and reading span scores. *Language and Cognition*, *6*(1), 111–145.

Hofmeister, P., Jaeger, T. F., Arnon, I., Sag, I. A., & Snider, N. (2013). The source ambiguity problem: Distinguishing the effects of grammar and processing on acceptability judgments. *Language and cognitive processes*, *28*(1-2), 48–87.

Hofmeister, P., & Sag, I. A. (2010). Cognitive constraints and island effects. *Language*, *86*(2), 366-415. Retrieved from `http://www.jstor.org/stable/40666324`

Ince, A. (2009). *Dimensions of ellipsis: Investigations in turkish* (Unpublished doctoral dissertation).

Ince, A. (2012). Sluicing in turkish. *Sluicing: Cross-linguistic perspectives (Oxford Studies in Theoretical Linguistics)*, *38*, 248-269.

Ioannidou, A., & Den Dikken, M. (2009). P-drop, d-drop, d-spread. In *Proceedings of the 2007 workshop in greek syntax and semantics* (pp. 393–408).

Jacobson, P. (2016). The short answer: implications for direct compositionality (and vice versa). *Language*, *92*(2), 331–375.

Johnson, K. (2001). *What vp ellipsis can do, and what it can't, but not why.* Citeseer.

Joseph, B. D., Philippaki-Warburton, I., & Philippaki-Warburton, I. (1987). *Modern greek.*

Croom Helm London.

Jurafsky, D. (1996). A probabilistic model of lexical and syntactic access and disambiguation. *Cognitive science*, *20*(2), 137–194.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: from eye fixations to comprehension. *Psychological review*, *87*(4), 329.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: individual differences in working memory. *Psychological review*, *99*(1), 122.

Just, M. A., Carpenter, P. A., & Woolley, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of experimental psychology: General*, *111*(2), 228.

Kaan, E., Overfelt, C., Tromp, D., & Wijnen, F. (2013). Processing gapped verbs. *Journal of psycholinguistic research*, *42*(4), 307–338.

Kaan, E., Wijnen, F., & Swaab, T. Y. (2004). Gapping: Electrophysiological evidence for immediate processing of 'missing' verbs in sentence comprehension. *Brain and Language*, *89*(3), 584–592.

Kaiser, E., Runner, J. T., Sussman, R. S., & Tanenhaus, M. K. (2005). What influences the referential properties of reflexives and pronouns in finnish?

Kayne, R. S. (1984). Connectedness and binary branching.

Kazanina, N., Lau, E. F., Lieberman, M., Yoshida, M., & Phillips, C. (2007). The effect of syntactic constraints on the processing of backwards anaphora. *Journal of Memory and Language*, *56*(3), 384–409.

Kehler, A. (2000). Coherence and the resolution of ellipsis. *Linguistics and Philosophy*, *23*(6), 533–575.

Kempen, G. (2010). Grammatical movement and gap-filling: Unproven cognitive processes. *Ms. Max Planck Institute for Psycholinguistics, Nijmegen & Leiden University*.

Kemper, S., Thompson, M., & Marquis, J. (2001). Longitudinal change in language production: effects of aging and dementia on grammatical complexity and propositional content. *Psychology and aging*, *16*(4), 600.

Keuleers, E., Brysbaert, M., & New, B. (2010). Subtlex-nl: A new measure for dutch word frequency based on film subtitles. *Behavior research methods*, *42*(3), 643–650.

Kim, C. S., Kobele, G. M., Runner, J. T., & Hale, J. T. (2011). The acceptability cline in vp ellipsis. *Syntax*, *14*(4), 318–354.

Klatte, M., Lachmann, T., Meis, M., et al. (2010). Effects of noise and reverberation on speech perception and listening comprehension of children and adults in a classroom-like setting. *Noise and Health*, *12*(49), 270.

Kleinschmidt, D. F., & Jaeger, T. F. (2015). Robust speech perception: recognize the familiar, generalize to the similar, and adapt to the novel. *Psychological review*, *122*(2), 148.

Komorita, S. S., & Graham, W. K. (1965). Number of scale points and the reliability of scales. *Educational and Psychological Measurement*.

Krosnick, J. A., & Fabrigar, L. R. (1997). Designing rating scales for effective measurement in surveys. *Survey measurement and process quality*, 141–164.

Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, cognition and neuroscience*, *31*(1), 32–59.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: finding meaning in the n400 component of the event-related brain potential (erp). *Annual review of psychology*, *62*, 621–647.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*(4427), 203–205.

Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature*, *307*(5947), 161.

Lakoff, G. (1972). The arbitrary basis of transformational grammar. *Language*, 76–87.

Lasnik, H. (2001). When can you save a structure by destroying it? In *Proceedings-nels* (Vol. 31, pp. 301–320).

Lasnik, H. (2005). Review of the syntax of silence, by jason merchant. *Language*, *81*, 259–265.

Lasnik, H. (2014). Multiple sluicing in english? *Syntax*, *17*(1), 1–20.

Lau, E., Stroud, C., Plesch, S., & Phillips, C. (2006). The role of structural prediction in rapid syntactic analysis. *Brain and language*, *98*(1), 74–88.

Lau, E. F., Holcomb, P. J., & Kuperberg, G. R. (2013). Dissociating n400 effects of prediction from association in single-word contexts. *Journal of cognitive neuroscience*, *25*(3), 484–502.

Law, P. (1998). A unified analysis of p-stranding in romance and germanic. In *Proceedings of nels* (Vol. 28, pp. 219–234).

Lefcheck, J. S. (2016). piecewisesem: Piecewise structural equation modeling in r for ecology, evolution, and systematics. *Methods in Ecology and Evolution*, *7*(5), 573-579. Retrieved from `http://dx.doi.org/10.1111/2041-210X.12512` doi: 10.1111/2041-210X.12512

Lemke, R., Schäfer, L., & Reich, I. (2018). Syntactic cues license voice mismatch in vp ellipsis – an experimental study. In *Talk at dgfs, stuttgart, 2018.*

Leonetti, M. (1999). El artículo. *Gramática descriptiva de la lengua española*, *1*, 787–890.

Leung, T. (2014). The preposition stranding generalization and conditions on sluicing: Evidence from emirati arabic. *Linguistic Inquiry*, *45*(2), 332–340.

Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions, and reversals. In *Soviet physics doklady* (Vol. 10, pp. 707–710).

Levin, L. (1982). Sluicing: A lexical interpretation procedure. *The mental representation of grammatical relations*, 590–654.

Levy, R. (2008a). Expectation-based syntactic comprehension. *Cognition*, *106*(3), 1126–1177.

Levy, R. (2008b). A noisy-channel model of rational human sentence comprehension under uncertain input. In *Proceedings of the conference on empirical methods in natural language processing* (pp. 234–243).

Luce, R. D. (2002). A psychophysical theory of intensity proportions, joint presentations, and matches. *Psychological Review*, *109*(3), 520.

MacDonald, M. C., Just, M. A., & Carpenter, P. A. (1992). Working memory constraints on the processing of syntactic ambiguity. *Cognitive psychology*, *24*(1), 56–98.

Magyari, L., & de Ruiter, J. P. (2012). Prediction of turn-ends based on anticipation of upcoming words. *Frontiers in psychology*, *3*, 376.

Manzini, M. R. (1998). A minimalist theory of weak islands. *SYNTAX AND SEMANTICS, VOL 29*, *29*, 185–209.

Marr, D. (1982). *Vision: A computational investigation into.* WH Freeman.

Marslen-Wilson, W. (1973). Linguistic structure and speech shadowing at very short latencies. *Nature*, *244*(5417), 522.

Marslen-Wilson, W., & Tyler, L. K. (1975). Processing structure of sentence perception. *Nature*, *257*(5529), 784.

Marslen-Wilson, W. D. (1975). Sentence perception as an interactive parallel process. *Science*, *189*(4198), 226–228.

Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive psychology*, *10*(1), 29–63.

Martin, A. E., & McElree, B. (2008). A content-addressable pointer mechanism underlies comprehension of verb-phrase ellipsis. *Journal of Memory and Language*, *58*(3), 879–906.

Martin, A. E., & McElree, B. (2011). Direct-access retrieval during sentence comprehension: evidence from sluicing. *Journal of memory and language*, *64*(4), 327–343.

Marušič, F., & Žaucer, R. (2013). A note on sluicing and island repair. In *Annual workshop on formal approaches to slavic linguistics, the third indiana meeting 2012* (pp. 176–189).

Masullo, P. J., & Depiante, M. (2004). Gender is in the lexicon, number is in the syntax: Evidence from nominal ellipsis in spanish. *Comunicación en GLOW*, *24*.

McCawley, J. D. (1988). Adverbial nps: bare or clad in see-through garb? *Language*, *64*(3), 583–590.

Merchant, J. (1998). Pseudosluicing: elliptical clefts in japanese and english. *ZAS working papers in linguistics*, *10*, 88–112.

Merchant, J. (2001). *The syntax of silence: Sluicing, islands, and the theory of ellipsis.* Oxford University Press.

Merchant, J. (2005a). Fragments and ellipsis. *Linguistics and philosophy*, *27*(6), 661–738.

Merchant, J. (2005b). Fragments and ellipsis. *Linguistics and philosophy*, *27*(6), 661–738.

Merchant, J. (2008). Variable island repair under ellipsis. *Topics in ellipsis*, 132–153.

Merchant, J. (2014). Gender mismatches under nominal ellipsis. *Lingua*, *151*, 9–32.

Mitchell, D. C. (1994). Sentence parsing. *Handbook of psycholinguistics*, 375–409.

Mitchell, D. C., Cuetos, F., Corley, M. M., & Brysbaert, M. (1995). Exposure-based models of human parsing: Evidence for the use of coarse-grained (nonlexical) statistical records. *Journal of Psycholinguistic Research*, *24*(6), 469–488.

Molimpakis, E. (2016a). Preposition-stranding under sluicing in modern greek. In *Ellipsis across borders conference.*

Molimpakis, E. (2016b). *Preposition-stranding under sluicing in modern greek.* Upgrade paper.

Monsell, S., Doyle, M., & Haggard, P. (1989). Effects of frequency on visual word recognition tasks: Where are they? *Journal of Experimental Psychology: General*, *118*(1), 43.

Morgan, J. (1973). Sentence fragments and the notion 'sentence'. *Issues in linguistics: Papers in honor of Henry and Renée Kahane*, 719–751.

Murray, W. S., & Forster, K. I. (2004). Serial mechanisms in lexical access: the rank hypothesis. *Psychological Review*, *111*(3), 721.

Myers, J. (2009). The design and analysis of small-scale syntactic judgment experiments. *Lingua*, *119*(3), 425–444.

Nairne, J. S. (1990). A feature model of immediate memory. *Memory & Cognition*, *18*(3), 251–269.

Nairne, J. S. (2001). A functional analysis of primary memory. *The nature of remembering: Essays in honor of Robert G. Crowder*, 282–296.

Nairne, J. S. (2006). Modeling distinctiveness: Implications for general memory theory. *Distinctiveness and memory*, 27–46.

Nakagawa, S., & Schielzeth, H. (2010). Repeatability for gaussian and non-gaussian data: a practical guide for biologists. *Biological Reviews*, *85*(4), 935–956.

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining r2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, *4*(2), 133–142.

Narayanan, S., & Jurafsky, D. (2002). A bayesian model predicts human parse preference and reading times in sentence processing. In *Advances in neural information processing systems* (pp. 59–65).

Narens, L. (1996). A theory of ratio magnitude estimation. *Journal of Mathematical Psychology*, *40*(2), 109–129.

Nykiel, J. (2013). Clefts and preposition omission under sluicing. *Lingua*, *123*, 74–117.

Nykiel, J. (2014). *The ellipsis alternation: remnants with and without prepositions.*

Manuscript.

Nykiel, J. (2015). Constraints on ellipsis alternation: A view from the history of english. *Language Variation and Change*, *27*(02), 227–254.

Oberauer, K., & Lewandowsky, S. (2008). Forgetting in immediate serial recall: Decay, temporal distinctiveness, or interference? *Psychological review*, *115*(3), 544.

Ott, D., & Therrien, R. (2018). *Swiping in a variety of ontario french.* Manuscript.

Pearlmutter, N. J., Garnsey, S. M., & Bock, K. (1999). Agreement processes in sentence comprehension. *Journal of Memory and language*, *41*(3), 427–456.

Phillips, C. (2006). The real-time status of island phenomena. *Language*, *82*(4), 795–823.

Phillips, C., & Lewis, S. (2013). Derivational order in syntax: Evidence and architectural consequences. *Studies in Linguistics*, *6*, 11–47.

Phillips, C., & Parker, D. (2014). The psycholinguistics of ellipsis. *Lingua*, *151*, 78–95.

Phillips, C., & Wagers, M. (2007). Relating structure and time in linguistics and psycholinguistics. *Oxford handbook of psycholinguistics*, 739–756.

Pickering, M., & Barry, G. (1991). Sentence processing without empty categories. *Language and Cognitive Processes*, *6*(3), 229-259. Retrieved from `http://dx.doi.org/10.1080/01690969108406944` doi: 10.1080/01690969108406944

Pollard, C., & Sag, I. A. (1992). Anaphors in english and the scope of binding theory. *Linguistic inquiry*, *23*(2), 261–303.

Pollard, C., & Sag, I. A. (1994). *Head-driven phrase structure grammar.* University of Chicago Press.

Pollmann, T. (1975). Een regel die subject en copula deleert. *Spektator*, *5*, 282–292.

Pritchett, B. L. (1988). Garden path phenomena and the grammatical basis of language processing. *Language*, 539–576.

Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). A comprehensive grammar of the english language. *London and New York: Longman*.

Rabbitt, P. M. (1968). Channel-capacity, intelligibility and immediate memory. *The Quarterly journal of experimental psychology*, *20*(3), 241–248.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological bulletin*, *124*(3), 372.

Rayner, K., Binder, K. S., Ashby, J., & Pollatsek, A. (2001). Eye movement control in reading: Word predictability has little influence on initial landing positions in words. *Vision Research*, *41*(7), 943–954.

Rayner, K., & Well, A. D. (1996). Effects of contextual constraint on eye movements in reading: A further examination. *Psychonomic Bulletin & Review*, *3*(4), 504–509.

Reichle, E. D., Reineberg, A. E., & Schooler, J. W. (2010). Eye movements during mindless reading. *Psychological Science*, *21*(9), 1300–1310.

Reinhart, T., & Reuland, E. (1993). Reflexivity. *Linguistic inquiry*, *24*(4), 657–720.

Richards, N. (1997). *What moves where when in which language? phd.* MIT, Cambridge, Mass.

Richards, N. (2001). *Movement in language: Interactions and architectures.* Oxford University Press, USA.

Rizzi, L. (1990). *Relativized minimality.* The MIT Press.

Rodrigues, C., Nevins, A., & Vicente, L. (2009). Cleaving the interactions between sluicing and p-stranding. *Romance languages and linguistic theory 2006: selected papers from 'Going Romance', Amsterdam, 7–9 December 2006*, 175–198.

Romero, M. (1998). *Focus and reconstruction effects in wh-phrases* (Unpublished doctoral dissertation). University of Massachusetts Amherst.

Rooth, M. (1992). A theory of focus interpretation. *Natural language semantics*, *1*(1), 75–116.

Rosen, C. (1976). Guess what about. In *Proceedings of nels* (Vol. 6, pp. 205–211).

Ross, J. R. (1967). Constraints on variables in syntax.

Ross, J. R. (1969). Guess who. *Sluicing: Cross-Linguistic Perspectives*.

Rottman, I., & Yoshida, M. (2013). Sluicing, idioms, and island repair. *Linguistic inquiry*, *44*(4), 651–668.

Runner, J. T., & Goldwater, M. B. (2011). On the representational nature of representational noun phrases. *Experiments at the Interfaces*, *37*, 219–250.

Runner, J. T., Sussman, R. S., & Tanenhaus, M. K. (2003). Assignment of reference to reflexives and pronouns in picture noun phrases: evidence from eye movements. *Cognition*, *89*(1), B1–B13.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1978). A simplest systematics for the organization of turn taking for conversation. In *Studies in the organization of conversational interaction* (pp. 7–55). Elsevier.

Sag, I. A. (1976). *Deletion and logical form.* (Unpublished doctoral dissertation). Massachusetts Institute of Technology.

Sag, I. A., & Nykiel, J. (2011). Remarks on sluicing. In *Proceedings of the 18th international conference on head-driven phrase structure grammar (hpsg 2011)* (pp. 188–208).

Sanford, A. J., Moar, K., & Garrod, S. C. (1988). Proper names as controllers of discourse focus. *Language and speech*, *31*(1), 43–56.

Sato, Y. (2011). P-stranding under sluicing and repair by ellipsis: why is indonesian (not) special? *Journal of East Asian Linguistics*, *20*(4), 339–382.

Sauerland, U. (1996). Guess how. In *Proceedings of console* (Vol. 4, pp. 297–309).

Sayette, M. A., Schooler, J. W., & Reichle, E. D. (2010). Out for a smoke: The impact of cigarette craving on zoning out during reading. *Psychological science*, *21*(1), 26–30.

Schad, D. J., Nuthmann, A., & Engbert, R. (2012). Your mind wanders weakly, your mind wanders deeply: objective measures reveal mindless reading at different levels.

Cognition, *125*(2), 179–194.

Schooler, J. W., Smallwood, J., Christoff, K., Handy, T. C., Reichle, E. D., & Sayette, M. A. (2011). Meta-awareness, perceptual decoupling and the wandering mind. *Trends in cognitive sciences*, *15*(7), 319–326.

Schwarzschild, R. (1999). Givenness, avoidf and other constraints on the placement of accent. *Natural language semantics*, *7*(2), 141–177.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell system technical journal*, *27*(3), 379–423.

Shapiro, L. P., Hestvik, A., Lesan, L., & Garcia, A. R. (2003). Charting the time-course of vp-ellipsis sentence comprehension: Evidence for an initial and independent structural analysis. *Journal of Memory and Language*, *49*(1), 1 - 19. Retrieved from `http://www.sciencedirect.com/science/article/pii/S0749596X03000263` doi: http://dx.doi.org/10.1016/S0749-596X(03)00026-3

Sloggett, S. (2013). Case licensing in processing: Evidence from german. In *Poster at the 26th annual cuny conference on human sentence processing. columbia, sc.*

Smith, N. J., & Levy, R. (2013). The effect of word predictability on reading time is logarithmic. *Cognition*, *128*(3), 302–319.

Snider, N., & Runner, J. T. (2010). Structural parallelism aids ellipsis and anaphor resolution: Evidence from eye movements to semantic and phonological neighbors. *Architectures and Mechanisms of Language Processing*.

Spivey-Knowlton, M. J., Trueswell, J. C., & Tanenhaus, M. K. (1993). Context effects in syntactic ambiguity resolution: Discourse and semantic influences in parsing reduced relative clauses. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, *47*(2), 276.

Sprouse, J. (2006). The accent projection principle: Why the hell not? *University of Pennsylvania Working Papers in Linguistics*, *12*(1), 28.

Sprouse, J. (2007). *A program for experimental syntax: Finding the relationship between acceptability and grammatical knowledge.* ProQuest.

Sprouse, J. (2008). The differential sensitivity of acceptability judgments to processing effects. *Linguistic Inquiry*, *39*(4), 686–694.

Sprouse, J. (2011). A test of the cognitive assumptions of magnitude estimation: Commutativity does not hold for acceptability judgments. *Language*, 274–288.

Sprouse, J., Wagers, M., & Phillips, C. (2012). A test of the relation between working-memory capacity and syntactic island effects. *Language*, *88*(1), 82–123.

Staub, A. (2015). The effect of lexical predictability on eye movements in reading: Critical review and theoretical interpretation. *Language and Linguistics Compass*, *9*(8), 311–327.

Staub, A., & Clifton Jr, C. (2006). Syntactic prediction in language comprehension: Evidence from either... or. *Journal of experimental psychology: Learning, memory, and cognition*, *32*(2), 425.

Staub, A., Grant, M., Astheimer, L., & Cohen, A. (2015). The influence of cloze probability and item constraint on cloze task response time. *Journal of Memory and Language*, *82*, 1–17.

Staum Casasanto, L., Hofmeister, P., & Sag, I. A. (2010). Understanding acceptability judgments: Additivity and working memory effects..

Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., ... others (2009). Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences*, pnas–0903616106.

Stjepanović, S. (2008). P-stranding under sluicing in a non-p-stranding language? *Linguistic Inquiry*, *39*(1), 179–190.

Stolcke, A. (1995). An efficient probabilistic context-free parsing algorithm that computes prefix probabilities. *Computational linguistics*, *21*(2), 165–201.

Sturt, P. (2003). The time-course of the application of binding constraints in reference resolution. *Journal of Memory and Language*, *48*(3), 542–562.

Sullivan, G. M., & Feinn, R. (2012). Using effect size—or why the p value is not enough. *Journal of graduate medical education*, *4*(3), 279–282.

Swaab, T. Y., Camblin, C. C., & Gordon, P. C. (2004). Electrophysiological evidence for reversed lexical repetition effects in language processing. *Journal of cognitive Neuroscience*, *16*(5), 715–726.

Szczegielniak, A. (2006). All sluiced up, but no alleviation in sight...

Tanenhaus, M. K., Carlson, G., & Trueswell, J. C. (1989). The role of thematic structures in interpretation and parsing. *Language and cognitive processes*, *4*(3-4), SI211–SI234.

Tanenhaus, M. K., & Carlson, G. N. (1990). Comprehension of deep and surface verbphrase anaphors. *Language and Cognitive Processes*, *5*(4), 257-280. Retrieved from `http://dx.doi.org/10.1080/01690969008407064` doi: 10.1080/01690969008407064

Tang, K. (2014–2018). *Linger toolkit.* `www.kevintang.org/Tools.html`. Github.

Taylor, W. L. (1953). "cloze procedure": A new tool for measuring readability. *Journalism Bulletin*, *30*(4), 415–433.

Tenenbaum, J. B. (1999). *A bayesian framework for concept learning* (Unpublished doctoral dissertation). Massachusetts Institute of Technology.

Thompson, C. K., Shapiro, L. P., & Roberts, M. M. (1993). Treatment of sentence production deficits in aphasia: A linguistic-specific approach to wh-interrogative training and generalization. *Aphasiology*, *7*(1), 111–133.

Traxler, M. J. (2014). Trends in syntactic parsing: Anticipation, bayesian estimation, and good-enough parsing. *Trends in cognitive sciences*, *18*(11), 605–611.

Traxler, M. J., Pickering, M. J., & Clifton Jr, C. (1998). Adjunct attachment is not a form of lexical ambiguity resolution. *Journal of Memory and Language*, *39*(4), 558–592.

Van Craenenbroeck, J. (2004). *Ellipsis in dutch dialects.* Netherlands Graduate School of Linguistics.

Van Craenenbroeck, J. (2009). What sluicing can do, what it can't and in which language-on the cross-linguistic syntax of ellipsis.

Van Craenenbroeck, J. (2010a). Invisible last resort: A note on clefts as the underlying source for sluicing. *Lingua*, *120*(7), 1714–1726.

Van Craenenbroeck, J. (2010b). *The syntax of ellipsis: Evidence from dutch dialects.* Oxford University Press.

Van Craenenbroeck, J. (2012). Ellipsis, identity, and accommodation. *Ms. Hogeschool-Universiteit Brussel*.

Van Craenenbroeck, J., & Lipták, A. (2006). The crosslinguistic syntax of sluicing: Evidence from hungarian relatives. *Syntax*, *9*(3), 248–274.

Van Craenenbroeck, J., & Lipták, A. (2013). What sluicing can do, what it can't and in which language: On the cross-linguistic syntax of ellipsis.

Van Gompel, R. P., & Liversedge, S. P. (2003). The influence of morphological information on cataphoric pronoun assignment. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(1), 128.

Van Gompel, R. P., Pickering, M. J., Pearson, J., & Liversedge, S. P. (2005). Evidence against competition during syntactic ambiguity resolution. *Journal of Memory and Language*, *52*(2), 284–307.

Van Gompel, R. P., Pickering, M. J., & Traxler, M. J. (2001). Reanalysis in sentence processing: Evidence against current constraint-based and two-stage models. *Journal of Memory and Language*, *45*(2), 225–258.

Van Heuven, W. J., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). Subtlex-uk: A new and improved word frequency database for british english. *The Quarterly Journal of Experimental Psychology*, *67*(6), 1176–1190.

Van Petten, C., & Kutas, M. (1990). Interactions between sentence context and word frequencyinevent-related brainpotentials. *Memory & Cognition*, *18*(4), 380–393.

Van Petten, C., & Luka, B. J. (2012). Prediction during language comprehension: Benefits, costs, and erp components. *International Journal of Psychophysiology*, *83*(2), 176–190.

Van Riemsdijk, H. (1998). Head movement and adjacency. *Natural Language & Linguistic Theory*, *16*(3), 633–679.

Vicente, L. (2015). Morphological case mismatches under sluicing. *in upcoming issue of Snippets*.

Vicente, L. (2018). Sluicing and its subtypes. *The Oxford handbook of ellipsis*.

Vitevitch, M. S., & Luce, P. A. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language*, *40*(3), 374–408.

Wagers, M. W., & Phillips, C. (2009). Multiple dependencies and the role of the grammar in real-time comprehension. *Journal of Linguistics*, *45*(2), 395–433.

Wang, C.-a. A. (2007). Sluicing and resumption. In E. Elfner & M. Walkow (Eds.), (Vol. 2, p. 239-252). Amherst: University of Massachusetts.

Wasow, T. (1972). *Anaphoric relations in english* (Unpublished doctoral dissertation). Massachusetts Institute of Technology Cambridge.

Weir, A. (2014). Fragments and clausal ellipsis: University of massachusetts amherst dissertation.

Weskott, T., & Fanselow, G. (2011). On the informativity of different measures of linguistic acceptability. *Language*, 249–273.

Westfall, J., Kenny, D. A., & Judd, C. M. (2014). Statistical power and optimal design in experiments in which samples of participants respond to samples of stimuli. *Journal of Experimental Psychology: General*, *143*(5), 2020.

Wiese, H. (2009). Grammatical innovation in multiethnic urban europe: New linguistic practices among adolescents. *Lingua*, *119*(5), 782–806.

Williams, E. S. (1977). Discourse and logical form. *Linguistic inquiry*, 101–139.

Wingfield, A., Tun, P. A., & McCoy, S. L. (2005). Hearing loss in older adulthood: What it is and how it interacts with cognitive performance. *Current directions in psychological science*, *14*(3), 144–148.

Wlotko, E. W., & Federmeier, K. D. (2012). So that's what you meant! event-related potentials reveal multiple aspects of context use during construction of message-level meaning. *NeuroImage*, *62*(1), 356–366.

Wood, J., Barros, M., & Sigurðsson, E. F. (2016). Clausal ellipsis and case (mis)matching in icelandic. *Working papers in Scandinavian Syntax*, *96*, 49-90.

Yoshida, M., Dickey, M. W., & Sturt, P. (2013). Predictive processing of syntactic structure: Sluicing and ellipsis in real-time sentence processing. *Language and Cognitive Processes*, *28*(3), 272-302. Retrieved from `http://dx.doi.org/10.1080/01690965.2011.622905` doi: 10.1080/01690965.2011.622905