

Copyright
by
Joshua Frank
2020

**The Dissertation Committee for Joshua Frank Certifies that this is the approved
version of the following Dissertation:**

**An Experimental Approach to Re complementation:
Evidence from Monolingual and Bilingual Spanish**

Committee:

Almeida Jacqueline Toribio, Supervisor

Alejandro Cuza

Barbara Bullock

Sandro Sessarego

**An Experimental Approach to Re complementation:
Evidence from Monolingual and Bilingual Spanish**

by

Joshua Frank

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August 2020

Dedication

A mis padres, Vivian and Dan, who in *different* ways have always inspired me.

Acknowledgements

During the course of this dissertation project, I have been fortunate to count on the support of many. Thank you to my mentors Almeida Jacqueline Toribio and Alejandro Cuza for their unconditional support and belief in me, as well as their steadfast direction throughout the course of this project. I am grateful to my two additional committee members, Barbara Bullock and Sandro Sessarego, as well as several colleagues past and present—Fernando Llanos, Adrian Riccelli, Crystal Marull, Jesse Abing and David Giancaspro—who all provided valuable feedback and assistance. I am also deeply appreciative of Belem López and Llama lab for affording me new collaborative research opportunities and lab management experience during my final year as a graduate student.

Throughout the duration of my PhD program, I spent time not only in the “live music capital of the world” but also New Brunswick, NJ, during a fellowship year at Rutgers, as well as NYC, where I helped to develop an online language learning program at Lingo Live. I’m particularly grateful to my Noise Revival Orchestra bandmates, my classmates at both UT Austin and Rutgers, as well as my former work family at Lingo Live. The camaraderie leaves a fond and indelible mark on this chapter of my life.

Thank you to my parents, to whom this dissertation project is dedicated, and my siblings—Aron, Joey and Miriam—who bring me so much joy. Lastly, I’m grateful to Jessie and Teddy for their patience and trust as I converted our only bedroom into an office and a classroom. Thank you for insisting on a daily routine, including regular exercise. This was a different time and I’m grateful to have experienced it with the two of you.

Abstract

An Experimental Approach to Recomplementation: Evidence from Monolingual and Bilingual Spanish

Joshua Frank, Ph.D.

The University of Texas at Austin, 2020

Supervisor: Almeida Jacqueline Toribio

This dissertation advances the study of recombination in Spanish (e.g., Villa-García, 2015), with three experimental studies that probe the representation and processing of the left periphery while addressing shortcomings in the field of syntax more generally. Recombination is the phenomenon whereby one or more left-dislocated phrases or circumstantial adjuncts intervene between a primary (C1) and secondary (C2) complementizer, e.g., *He said that₁ later in the afternoon that₂ he would clean his room.* Study 1 investigates the grammatical status of recombination in US heritage speakers of Spanish via acceptability judgment and preference tasks. Results demonstrate that heritage speakers prefer the overt C2 variety at a higher rate than the baseline group. These findings are interpreted within the Model of Divergent Attainment (Polinsky & Scontras, 2020), where complexities associated with “silent” phenomena and dependency distance, along with processing burden, lead to reanalysis and eventual divergent attainment. Study 2 explores recombination as a locus of dialectal variation in Colombian and Cuban Spanish via elicited imitation and sentence completion tasks. Results provide evidence that overt C2 is neither licensed by the grammar nor a facilitator of complement integration.

Importantly, the possibility of task effect cannot be ruled out. Lastly, study 3 analyzes the incremental processing of recomplementation via self-paced reading. Results demonstrate that a psycholinguistic model informed by syntactic theory is favorable to one that is not. This conclusion is further supported by an analysis of individual differences in working memory span. While advancing recomplementation research, this dissertation offers experimental evidence in support of three broader claims. First, speakers with diverse profiles (e.g., heritage speakers) inform general theory and contribute to such disparate topics as processing complexity, the role of input and experience in language development and variation among the Spanishes of the world. Second, researcher selection bias and the effects of task must not be overlooked in the literature, as they threaten the ultimate pursuit of knowledge. Finally, when experimental findings, psycholinguistic models and syntactic-theoretical accounts inform one another, the outcome is superior.

Table of Contents

List of Tables	xii
List of Figures	xiv
Chapter 1: Introduction.....	1
1.1 Background.....	3
1.2 Re complementation	5
1.3 Specific Findings and Broader Contributions	7
1.4 Organization of the Dissertation.....	10
Chapter 2: US Heritage Spanish Study.....	12
2.1 Introduction	12
2.2 Linguistic Phenomenon.....	14
2.2.1 Re complementation Theory.....	14
2.2.2 Experimental Evidence.....	16
2.3 Sources of Divergent Attainment	20
2.3.1 Vulnerable Phenomena.....	20
2.3.2 Model of Divergent Attainment.....	22
2.3.3 Research Questions	26
2.4 The Study	28
2.4.1 Participants.....	28
2.4.2 Methods and Design.....	31
2.4.3 Results.....	35
2.5 Discussion	42

Chapter 3: Cuban and Colombian Spanish Study	49
3.1 Introduction	49
3.2 Linguistic Phenomenon.....	52
3.2.1 Syntactic-Theoretical Review	52
3.2.2 The Function of Secondary <i>Que</i>	57
3.2.3 The Disparity between Theoretical Accounts and Experimental Findings	58
3.3 Morphosyntactic Variation in Spanish.....	60
3.3.1 Cuban and “Mainland” Spanish	60
3.3.2 Frank and Toribio (2017).....	64
3.3.3 Research Questions	66
3.4 The Study	69
3.4.1 Participants.....	69
3.4.2 Methods and Design	71
3.4.3 Results.....	75
3.5 Discussion	84
Chapter 4: Incremental Sentence Processing Study	89
4.1 Introduction	89
4.2 Syntactic Parsing.....	92
4.2.1 Expectation-Based and Memory-Based Accounts	92
4.2.2 Committing to a Prediction.....	96
4.2.3 Working Memory	99
4.3 Competing Accounts of Re complementation.....	102
4.3.1 Memory-Based Accounts	102

4.3.2 An Expectation-Based Account	106
4.3.3 Research Questions	108
4.4 The Study	111
4.4.1 Participants.....	111
4.4.2 Experimental Methodology	113
4.5 Results	121
4.5.1 Pre-Analysis	121
4.5.2 Matrix Verb Analysis	122
4.5.3 Complement Integration Analysis.....	123
4.5.4 Working Memory Analysis.....	130
4.6 Discussion	133
Chapter 5: Conclusion	141
5.1 Summary	141
5.2 Limitations.....	149
5.3 Future Research	151
5.4 Final Remarks.....	154
Appendices.....	155
Appendix A: Linguistic Questionnaire and Self-Assessment.....	155
Appendix B: Written Spanish Proficiency Test (DELE).....	161
Appendix C: Stimuli for Acceptability Judgment Task	166
Appendix D: Stimuli for Preference Task	170
Appendix E: Stimuli for Elicited Imitation Task	174
Appendix F: Stimuli for Sentence Completion Task	178

Appendix G: Sample Stimuli for Reverse Digit Span Task	183
Appendix H: Master List for Self-Paced Reading Task	184
References.....	189

List of Tables

Table 2.1. Participant profile.....	30
Table 2.2. Test group proportion Spanish language use.....	31
Table 2.3. Mean acceptability ratings from 1 “totally acceptable” to 7 “totally unacceptable”	36
Table 2.4. Participants who favored null or overt C2 by a low/high amount.....	38
Table 2.5. Proportion overt C2 preference	39
Table 2.6. Participants who preferred null or overt C2 overall.....	41
Table 3.1. Selected demographic information	69
Table 3.2. Language history, use and contact.....	70
Table 3.3. Variables under investigation	72
Table 3.4. Elicited imitation descriptive results for Cuban Spanish	76
Table 3.5. Elicited imitation descriptive results for Colombian Spanish	76
Table 3.6. Elicited imitation between groups comparison.....	78
Table 3.7. Elicited imitation statistical analysis.....	79
Table 3.8. Trimodal distribution of age and overt C2 repetition accuracy.....	80
Table 3.9. Sentence completion descriptive results for Cuban Spanish.....	82
Table 3.10. Sentence completion descriptive results for Colombian Spanish.....	83
Table 3.11. Sentence completion task, between groups comparison	83
Table 4.1. Selected participant metadata	112
Table 4.2. Eight experimental conditions.....	116
Table 4.3. Three sections for each test item.....	117
Table 4.4. Content-word metrics.....	119
Table 4.5. Latin-square design	120
Table 4.6. Comprehension question results	121

Table 4.7. Region of interest for matrix verb analysis	122
Table 4.8. Descriptive statistics for matrix verb analysis.....	122
Table 4.9. Region of interest in <i>wh</i> -complement analysis.....	124
Table 4.10. Descriptive statistics for <i>wh</i> -complement integration analysis	124
Table 4.11. Summary of linear mixed effects analysis	126
Table 4.12. Pairwise comparison post hoc analysis	128
Table 4.13. A comparison of results, Casasanto and Sag and the present study	129
Table 4.14. Summary of correlations with working memory.....	131
Table 4.15. Evidence for memory-based versus expectation-based accounts	135

List of Figures

Figure 2.1. Baseline group mean acceptability ratings.....	35
Figure 2.2. Test group mean acceptability ratings	35
Figure 2.3. Proportion overt C2 preference	39
Figure 4.1. Casasanto and Sag’s (2008) self-paced reading results.....	102
Figure 4.2. Interaction predicted by DLT	103
Figure 4.3. Interaction predicted by DLT+ (Casasanto & Sag, 2008)	103
Figure 4.4. Alternative interpretation of the interaction.....	104
Figure 4.5. Moving window paradigm with word-by-word segmentation.....	114
Figure 4.6. Present study self-paced reading results	127
Figure 4.7. Casasanto and Sag self-paced reading results	127
Figure 4.8. Negative correlation between working memory and overt/long condition ..	132
Figure 4.9. Positive correlation between working memory and <i>que</i> region, short-long conditions.....	132
Figure 4.10. An expectation-based model of recomplementation	138

Chapter 1: Introduction

Linguists observe language in natural and experimental environments, test hypotheses and propose theories or models. While the pursuit of linguistic knowledge is never complete or perfect, certain applications of scientific inquiry suffer more than others. With respect to the field of syntax, observations are collected in a number of different ways, ranging from introspections to formal experiments. These two modes of operation need one another as their limitations are well documented, often with contradictory/irreconcilable findings (e.g., Aarts, 2015; Collins, 2016; Edelman & Christiansen 2003; Ferreira, 2005; Gibson & Fedorenko 2010ab; Ortega-Santos, 2020ab). Selection bias, questions around methods and methodologies, as well as a disparity between syntactic-theoretical accounts and experimental findings continue to linger. The extant literature on recomplementation serves as a fine example. Recomplementation is the phenomenon whereby one or more left-dislocated (LD) phrases or circumstantial adjuncts are sandwiched between a high/primary complementizer (C1) and a low/secondary (C2) complementizer, as in the Spanish example in (1.1) and (1.2):

- (1.1) Me dice que₁ por suerte, que₂ va a tener suficiente tiempo.
CL₁SgDat say₃SgPres that for luck that is₃SgPres going to have enough time
'S/he says that luckily s/he is going to have enough time.'
- (1.2) Me pregunta que₁ esa camisa, que₂ cuándo la iba a devolver.
CL₁SgDat ask₃SgPres that that shirt that when CL₃SgFemAcc is₃SgPres going to return
'S/he asks when I am going to return the shirt.'

Utterances with multiple complementizers are attested in an array of embedded statement and question constructions. Further, they are not exclusive to Spanish. The phenomenon is also documented in languages such as Portuguese (e.g., Mascarenhas, 2007), Catalan (González i Planas, 2014), Galician (e.g., Gupton 2010), Italian (e.g.,

Ledgeway, 2000; Paoli, 2006; Rizzi, 2013) and English (e.g., Radford, 2018, 2013). Rigorous formal debates are ongoing (for a review see Villa-García, 2015), and experimental (e.g., Casasanto & Sag, 2018; Frank, 2016; Frank & Toribio, 2017) and corpus-based (Echeverría & López Seoane, 2019) studies provide evidence in favor of or against theories of representation and processing. Importantly, extant studies of recomplementation are limited by an insufficient representation of understudied dialects (e.g., the Caribbean lect) and speakers with diverse profiles (e.g., heritage speakers). Further, questions remain around the effects of certain experimental tasks on results, as well as the lack of attention to microvariations and individual speaker differences. Lastly, some syntactic-theoretical accounts and experimental findings are difficult to reconcile, forcing researchers to pick sides.

The primary goal of this dissertation is to offer three experimental studies on recomplementation to account for limited data in the areas of bilingualism, dialectal variation and syntactically informed psycholinguistic models. The topics of linguistic population biases, methodologies and methods, as well as the gap between syntactic-theoretical accounts and experimental findings are considered throughout. Specifically, study 1 investigates the grammars of bilingual heritage speakers of Spanish via acceptability judgment and preference tasks; study 2 explores dialectal variation (i.e., Colombian versus Caribbean) via elicited imitation and sentence completion tasks; and study 3 analyzes online comprehension via self-paced reading, while investigating individual differences in working memory as a predictor of performance. While these three studies adopt recomplementation as the linguistic phenomenon in question, this dissertation provides a broadly applicable roadmap to account for limitations and gaps that are all too common in the field.

1.1 BACKGROUND

Syntactic theory and the evidence that informs it is often limited in scope. The vast majority of studies adopt monolinguals as the exemplar speaker through which theoretical accounts are formed (e.g., Birdsong & Gertken, 2013). Crucially, native speaking bilingual populations have been shown to diverge from relevant baselines across several grammatical properties, including subject-verb inversion (e.g., Cuza, 2013; Cuza & Pérez-Tatam, 2016), number and gender agreement (e.g., Scontras, Polinsky & Fuchs, 2018), *pro*-drop (e.g., Montrul, 2002, 2004, 2008, 2016), mood selection (e.g., Giancaspro, 2017; Perez-Cortes, 2016) and clitic expression (e.g., Cuza, 2013; Montrul, 2010). Further, left periphery-related phenomena have been argued to be particularly prone to divergence (e.g., e.g., Bruhn de Garavito, 2002; Cuza, 2013; Frank, 2013; Isabelli, 2004; Montrul, 2010; Silva-Corvalán, 1993; Sorace, 2000, 2005). Even within a given linguistic population, bilingual or monolingual, the outcome of first language acquisition has been argued to be variable. Different speakers can be understood to have different grammatical representations (e.g., Dabrowska, 2012). This variation is thought to be driven by factors pertaining to linguistic experience (e.g., quantity and quality of input) and individual cognitive traits (e.g., working memory, motivation). In summary it is important for theoretical and experimental accounts alike to investigate different linguistic populations, as well as extend the analysis to the individual social and cognitive level.

Research has also been heavily biased towards Indo-European languages, and then, only to selected varieties. With reference to Spanish specifically, Peninsular has been held as the exemplar variety. This is problematic given the variation across the Spanish dialects of the world, from Peninsular to Latin American and Caribbean to non-Caribbean. According to Ortega-Santos (2020a), “...Latin American Spanish is different from American English in that it evolved in a multinational context, with intensive language-

contact with indigenous languages” (12). Thus, we might expect greater microvariation in the former. Further, studies that have investigated convergence rates between formal and informal methods in Spanish and English have found lower rates of convergence in the former (~75%) than the latter (93%) (Ortega-Santos, 2020a; Sprouse et al., 2013). For example, the investigation of morphosyntactic properties in lesser-studied Caribbean populations has informed long held generalizations about Spanish. We now know that traditional assumptions related to pronominal expression and free variation of subject-verb word orders had carried a Peninsular or ‘Mainland’ bias (e.g., Camacho, 2013; Cuza, 2017; Lipski, 1997; Martinez-Sanz, 2011; Orozco, 2015; Ortiz López, 2010; Otheguy & Zentella, 2012; Toribio, 2000).

The gap between syntactic-theoretical accounts and experimental findings is in part rooted in philosophical outlook. Some researchers adopt introspective approaches to data collection in the service of theory, while others are more “data first” and design formal experiments to surface facts about language. Formal experiments are particularly valuable because they are efficient and allow the investigator to analyze less frequent forms, even impossible forms, as well as facilitate distinction between performance errors and facts (e.g., Ludlow, 2011). Importantly, both camps use data and observation as a window into a participant’s underlying representations. An elicited imitation or repetition task, for example, assumes constructions that are consistent with the speakers’ grammars will be reproduced accurately whereas those that are not will be altered (e.g., Hamayan et al., 1978). An acceptability judgment task, either formal or informal, assumes that a decrement in acceptability judgment is a devaluation of grammatical status. Methodological limitations across both paradigms can have a profound effect on the results and interpretation of data. Syntactic-theoretical accounts can suffer from researcher intuition or selection bias (Edelman & Christiansen 2003; Ferreira, 2005; Gibson & Fedorenko

2010ab; Ortega-Santos, 2020ab) and experimental findings are subject to a replication crisis (e.g., Aarts, 2015; Collins, 2016; Ortega-Santos, 2020ab).

In summary, tensions between syntactic-theoretical accounts and experimental findings are beyond surface deep. However, rather than take sides, experimental and theoretical linguistic camps must continue to inform one another, while challenging themselves with a wide array of testing methods (e.g., online vs. offline), modes (e.g., oral vs. written) and replications. Crucially, without taking such elements as bilingualism, dialectal variation, individual differences and methodological effects into account, the pursuit of truth will remain elusive.

1.2 RECOMPLEMENTATION

The syntax of recomplementation is a topic of frequent analysis among theoreticians (e.g., Brovotto, 2002; Demonte & Fernández-Soriano, 2009; Fontana, 1993; González i Planas, 2014; Gupton, 2010; Kempchinsky, 2013; Ledgeway, 2000; López, 2009; Martín-González, 2002; Mascarenhas, 2007; Paoli, 2006; Radford, 2013, 2018; Rodríguez-Ramalle, 2003; Uriagereka, 1995; Villa-García, 2012, 2015, 2019). This is in large part because the phenomenon facilitates a broader debate on the cartography of the left periphery (e.g., Rizzi, 1997, 2013) in Generativist and Minimalist programs (Chomsky 1965 et seq.). Varying theoretical accounts of recomplementation have been put forth. These proposals range from pre-split CP accounts (e.g., CP Recursion, see Fontana, 1993) to accounts that factor in functional projections to varying degrees (e.g., FinitenessP, see Brovotto, 2002; Discourse Projection, see Kempchinsky, 2013; Moving Complementizers, see Ledgeway, 2000; DoubledForceP, see Martín-González, 2002; and TopicP, see Rodríguez-Ramalle, 2003). While syntactic-theoretical debates on recomplementation

have not yet been resolved, they have converged on the grammaticality of the phenomenon and the optionality of C2 lexicalization (i.e., overt secondary *que*).

The precise function or purpose of C2 lexicalization also continues to be a topic of discussion. Functional accounts range from reintroducing force (e.g., Martín-González, 2002), marking topic (e.g., Rodríguez-Ramalle, 2003), serving as a processing marker (e.g., Casasanto & Sag, 2008) or marking discourse (e.g., Villa-García, 2019). Alternatively, Villa-García (2019) argues that optional secondary *que* is multifunctional and can serve multiple roles at once. In some instances, syntactic-theoretical accounts and functional accounts are closely related. For example, the DoubledForceP proposal (e.g., Martín-González, 2002) argues that high and low complementizers merge in ForceP and doubled-ForceP, respectively. The logic being that the lexicalized C2 reintroduces the complement and force of the sentence. The TopicP analysis, on the other hand, holds that the low complementizer heads TopicP, while the sandwiched phrase merges in the specifier position of that same phrase (e.g., Rodríguez-Ramalle, 2003; Villa-García, 2012, 2015). Lexicalized C2 then serves as a topic marker, referring to entities previously mentioned in the discourse.

While both DoubledForceP and TopicP accounts represent a unification between the syntactic-theoretical and the functional, they also speak to a broader tension in the literature. The former account argues that secondary *que* does not introduce any new information, while the latter argues that it does. This seemingly minor difference reverberates beyond the field of syntax. A growing body of experimental research argues that the underlying assumption of the grammatical status of secondary *que* deserves further scrutiny (Casasanto & Sag, 2008; Frank, 2016; Frank & Toribio, 2017). Casasanto and Sag (2008) argue that C2 lexicalization is associated with a grammatical violation that is overridden by the benefit it brings to real-time sentence processing. Following the logic of

the DoubledForceP account, they argue that C2 is a reiteration of C1 that provides no new information. Further, Frank (2016) provides experimental evidence that overt C2 is associated with a decrement in acceptability when compared to the null counterpart and Frank and Toribio (2017) show that overt C2 items are reproduced with less accuracy in elicited imitation.

In summary, the extant literature across syntactic-theoretical accounts and experimental or quantitative findings proceeds apace: theoretical accounts continue to debate competing abstract representations; the one or many functions of the secondary *que* are under consideration; and the assumption of grammaticality is scrutinized in experimental findings and processing-based accounts. Crucially, there are two threads that unify this disparate literature. First, researchers agree that recomplementation exists in present-day Spanish. Second, evidence across the paradigms converges on the relationship between C2 lexicalization and intervener length (e.g., Casasanto & Sag, 2008; Echeverría & López Seoane, 2019; Ledgeway, 2000; Radford, 2018; Villa-García, 2019). Specifically, secondary *que* is a complexity correlate, where it is more likely to be spelled out when intervener length is long (or complex). In slightly different terms, C2 lexicalization is probabilistically predicted as a function of distance/complexity.

1.3 SPECIFIC FINDINGS AND BROADER CONTRIBUTIONS

With the general background information and specific recomplementation literature in mind, this dissertation introduces three studies that contribute to bilingualism, dialectal variation and processing theory. While this dissertation is narrow in its focus on recomplementation, it is broad in scope and generalizable to the field as the whole. That is to say, the following three studies provide a blueprint that can be broadly applied for the

advancement of linguistic research. The three primary areas of questioning that drive this thesis are as follows:

- (RQ1) Do advanced heritage speaker grammars diverge from a relevant baseline? If so, how does bilingual data inform syntactic-theoretical accounts of recomplementation?
- (RQ2) Is recomplementation a locus of dialectal variation? If so, does this help to explain the divide in theoretical accounts and experimental findings?
- (RQ3) How can we reconcile psycholinguistic models of recomplementation with syntactic-theoretical accounts? Is it a worthy pursuit?

(RQ1) is specifically considered in the first study. This study, Frank (submitted), is designed to investigate divergent outcomes in heritage language competence by offering novel data in support of a predictive model of heritage grammar. Results from two experimental tasks show heritage speakers prefer the overt C2 variety at a higher rate than the baseline. In line with Polinsky and Scontras (2020), this study argues that silence and distance, along with processing burden, lead to a reanalysis of the linguistic phenomenon and eventual divergent attainment. Processing literature argues that the interpretation of missing elements and retrieval of information at a distance increases complexity (e.g., Gibson, 1998). This strain should be even greater in heritage speakers, who are holding multiple languages in parallel, while working in the less dominant one. The broader contribution of this study pertains to how speakers from diverse profiles can inform general theory. Heritage speaker populations provide a window into how underlying representations can differ across and within linguistic populations. They offer new perspectives on the nature of grammatical representations and linguistic complexity.

(RQ2) is explored in the second study. This study counters the Peninsular bias in the syntactic-theoretical literature by investigating dialects of Spanish in Havana, Cuba and Medellín, Colombia. These two varieties were selected since the Caribbean lect shares

morphosyntactic properties that differ from ‘mainland’ or Peninsular varieties. Results from two formal production tasks suggest that the secondary *que* neither facilitates complement integration nor is it licensed by the grammar. This finding is robust across fixed effects of intervener length (i.e., two words and seven words), intervener type (i.e., object pronoun and adverbial) and sentence type (i.e., question and statement). In summary, there is a reinforcement of the gap between the syntactic-theoretical literature which argues for the grammatical status of C2 lexicalization (e.g., Demonte & Fernández-Soriano, 2009; Villa-García, 2015) and experimental findings which provide evidence that C2 lexicalization is not licensed by the grammar (e.g., Casasanto & Sag, 2008; Frank, 2016; Frank & Toribio, 2017). Broader contributions pertain to the value of investigating understudied dialects and linguistic microvariations. For example, incongruencies between theoretical and experimental findings can be more deeply scrutinized. In the case of the present study, we move from a skepticism towards the assumptions made in theoretical accounts to questions pertaining to the effect of experimental method. The binary accurate/inaccurate dependent measure of the elicited imitation task may not be sufficiently sensitive to investigate recomplementation. We also note a potential conflation of production and comprehension effects. Lastly, a more direct measure of facilitation effects would adopt online methods.

The final area of questioning, (RQ3), is investigated in study 3. This study explores the incremental processing of recomplementation via online methodology. The investigation begins by scrutinizing Casasanto and Sag’s (2008) memory-based model, which is incompatible with the syntactic-theoretical literature. Their model stipulates that overt C2 is the ungrammatical reiteration of C1. According to the authors, the phenomenon exists as part of a processing/grammar tradeoff when dependency distance is long. We propose an alternative expectation-based account that is informed by the syntactic-

theoretical literature. Novel data via self-paced reading paradigm is offered to determine whether a memory-based or an expectation-based model better accounts for the phenomenon. Results, at the group and individual level largely support the expectation-based account. The broader contribution of this study speaks to the importance of proposing psycholinguistic models that are grounded in linguistic theory. In the case of the present study, by developing a theory-informed model, the outcome was superior to the alternative. Further, the study argues for the value of including an analysis at the level of individual differences (e.g., working memory capacity).

1.4 ORGANIZATION OF THE DISSERTATION

The proceeding chapters are divided into three experimental studies and an overarching conclusion. Chapter 2 pertains to the study on bilingualism, specifically investigating the US heritage Spanish population within the model of divergent attainment (Polinsky & Scontras, 2020). After a brief introduction, motivating the study and introducing primary goals, literature reviews on recomplementation and divergent attainment in US heritage Spanish are offered. The remainder of the chapter pertains to the study itself, including research questions, participants, methods, results and discussion. Chapter 3 pertains to the study on dialectal variation, specifically building on Frank and Toribio (2017) by investigating both Cuban and Colombian Spanish. A short introduction, which also motivates the study and introduces primary goals, is preceded by literature reviews on recomplementation and morphosyntactic variation in Spanish. As in the previous chapter, the remainder of the study pertains to the investigation itself. The final investigation, chapter 4, pertains to the study on sentence processing, which scrutinizes Casasanto and Sag (2008), comparing memory-based versus expectation-based models of

recomplementation. After an introduction, relevant literature reviews on syntactic parsing and competing online accounts of recomplementation are discussed. What follows is the study itself. The dissertation concludes with Chapter 5, which synthesizes the conclusions of the three studies in order to highlight the primary contributions of the overall project, as well as the potential areas for future investigation.

Chapter 2: US Heritage Spanish Study

2.1 INTRODUCTION

In the keynote article for *Bilingualism: Language and Cognition* (Cambridge), Polinsky and Scontras (2020) lay out a preliminary roadmap for modeling the divergent morphosyntactic properties of heritage languages. Importantly, this divergent attainment model is meant to serve as a framework for predicting and discussing the differences in language competence between heritage speakers and a relevant baseline group (i.e., ideally the source of the heritage speakers' input); it is not intended to facilitate arguments around persistent difficulty or incomplete acquisition. In the framework, the authors propose “problematic” areas of language that can be viewed as sources of divergence in heritage grammar:

- (i) Morphology (e.g., agreement and marking)
- (ii) Distance (i.e., dependency relations at a distance)
- (iii) Silence (i.e., the interpretation of null elements)
- (iv) Ambiguity (i.e., one-to-many mappings between form and meaning)

This is not an exhaustive list, as the authors acknowledge. However, these four sources offer a novel way to categorize the literature on divergent attainment in heritage language acquisition and also help to focus the impending research program. The authors submit that the roadmap can “...serve as a jumping-off point for further progress toward a model. In particular, it can lead to specific empirical predictions about the ways in which heritage languages will (and will not deviate) from their respective baselines” (13).

With this framework in mind, the present study investigates the knowledge that advanced US heritage speakers of Spanish have of recomplementation. As discussed in the opening chapter, this CP-related phenomenon comprises a dislocated argument or circumstantial adjunct sandwiched between two complementizers. The argument or

adjunct is said to establish a specifier-head relationship in TopicP with secondary *que* ‘that’ (e.g., Villa-García, 2019), as exemplified by (1.1), reproduced here as (2.1):

(2.1) Me dice que₁ por suerte que₂ va a tener suficiente tiempo.
CL_{1SG} says that for luck that is going to have enough time
‘S/he says that luckily s/he is going to have enough time.’

While recomplementation is a topic of frequent investigation in formal circles (e.g., BrovETTO, 2002; Demonte & Fernández-Soriano, 2009; Fontana, 1993; González i Planas, 2014; Gupton, 2010; Martín-González, 2002; Mascarenhas, 2007; Paoli, 2006; Radford, 2018, 2013; Rizzi, 2013; Villa-García, 2019, 2015, 2012), very few studies have investigated the phenomenon via experimental methods (e.g., Casasanto & Sag, 2008; Echeverría & López Seoane, 2019; Frank, 2016). In addition, despite claims of widespread use among the dialects of present-day Spanish (e.g., Demonte & Fernández-Soriano, 2009), no study to date has examined the acquisition of recomplementation constructions in heritage language populations. This is surprising given the growing body of literature that examines whether advanced heritage speakers pattern monolingual norms with respect to left-periphery related phenomena like recomplementation, e.g., verb-second, embedded clauses and *wh*-questions (e.g., Bruhn de Garavito, 2002; Cuza, 2013; Cuza & Frank, 2011, 2015; Frank, 2013; Montrul, 2010; Silva-Corvalán, 1993; Zapata et al., 2005).

The present study makes three primary contributions to the field of heritage language acquisition:

- (i) An initial investigation of the acquisition of recomplementation structures in a heritage language population.
- (ii) Novel data in support of the growing literature arguing that the left periphery is a “vulnerable” domain in heritage language competence when compared to a relevant baseline group.
- (iii) Evidence in support of the Model of Divergent Attainment in heritage grammar.

In pursuing these three contributions, the work also sheds light on how the investigation of lesser studied populations with diverse profiles can and should inform existing theoretical accounts. The study begins with an overview of the linguistic phenomenon and a discussion of divergent attainment in heritage language acquisition.

2.2 LINGUISTIC PHENOMENON

2.2.1 Re complementation Theory

As reviewed earlier, re complementation, complementizer doubling, or multiple or reduplicated complementizers is defined as one or more left-dislocated (LD) phrases sandwiched between a high/primary complementizer (C1) and a low/secondary (C2) complementizer. As noted in Villa-García (2019), “in Spanish, any phrase susceptible of being left-dislocated can be featured in re complementation environments” (2). The phenomenon is not exclusive to Spanish (e.g., it is attested in Portuguese, Catalan, Galician, Italian, and English) and is argued to be widespread and prevalent among the dialects of present-day Spanish (Brovetto, 2002; Casasanto & Sag, 2008; Demonte & Fernández-Soriano, 2009; Fontana, 1993; González i Planas, 2014; Gupton, 2010; Martín-González, 2002; Mascarenhas, 2007; Paoli, 2006; Radford, 2013; Rizzi, 2013; Villa-García, 2019, 2015, 2012). Two experimental examples of re complementation are represented by (2.2) and (2.3):

(2.2) Me dijo/preguntó que esa guitarra vieja (que) cuándo la iba a tocar.
 CL₁SgDat asked₃SgPast that that guitar old that when CL₃SgFemAcc is₃SgPres
 going play
 ‘S/He asked me when I was going to play that old guitar.’

(2.3) Me dijo que esa guitarra vieja (que) la iba a tocar.
 CL₁SgDat said₃SgPast that that guitar old that CL₃SgFemAcc is₃SgPres going to play
 ‘S/He told me that s/he was going to play that old guitar.’

A critical difference between examples (2.2) and (2.3) is that the former is an embedded question, while the latter is a statement. In (2.2), someone didn't know when I was going to play a specific guitar so they asked me. In (2.3), someone simply told me they were going to play the guitar.

While a more comprehensive review of constructions like (2.2) and (2.3) is presented in chapter 3, here we review the doubled-ForceP and TopicP proposals. The doubled-ForceP analysis proposes that high and low complementizers merge in ForceP and doubled-ForceP, respectively (Demonte & Fernández-Soriano, 2009, 2014; Martín-González, 2002). The logic being that the lexicalized C2 (i.e., overt C2) reintroduces the complement and force or semantic function of the sentence. Because of the dependence of C2 on C1, it stands to reason that the two elements merge in related projections. Furthermore, the left-dislocated phrase merges in TopicP, and the interrogative phrase lands in FocusP, as in Rizzi (1997). See example (2.4) below:

(2.4) [_{ForceP} [_{Force'} que [_{TopicP} esa guitarra vieja [_{Top'} [_{(Doubled)ForceP} [_{(Doubled)Force'} que [_{FocusP} cuándo [_{FinitenessP} [_{Fin'} ...]]]]]]]]]]]

The strength of the doubled-ForceP analysis is evident when one considers the multifunctional nature of the secondary complementizer. Villa-García (2019) summarizes the account by stating the secondary complementizer can function as a discourse marker, a topic marker, a reportative/quotative marker, a force marker and a processing marker (see Section 3.2.2). The doubled-ForceP analysis nicely demonstrates how the complementizer can function as a subordinator, serving as a reminder of the force of the embedded clause. One potential shortcoming of this analysis is that it does not merge the low complementizer and the left-dislocated, sandwiched element in the same phrase. This could be interpreted as an unwelcome outcome, given the contingency of the low complementizer on the left-dislocated element. Note that *Me dijo que que* 'He said that that' is not grammatical.

Building on previous accounts, Villa-García (2012, 2015, 2019) advances the Topic account illustrated in (2.5) below:

(2.5) [_{ForceP} [_{Force'} que [_{TopicP} esa guitarra vieja [_{Top'} que ... [_{FocusP} cuándo [_{FinitenessP} [_{Fin'...}]]]]]]]]

The Topic analysis holds that instead of merging in a doubled-ForceP, the low complementizer heads TopicP, while the sandwiched phrase merges in the specifier position of that same phrase. This analysis, like doubled-Force, also speaks to the multifunctionality of the secondary complementizer. Namely, it is proposed to serve as a topic-marker, referring to entities previously mentioned in the discourse (e.g., *la guitarra* ‘the guitar’). Importantly, while this formal review has focused exclusively on recomplementation in Spanish, Villa-García (2019) argues for a single, unified syntactic-theoretical account of the phenomenon for Spanish and English, citing a number of syntactic and distributional similarities across the two languages.

2.2.2 Experimental Evidence

While formal syntactic debates on recomplementation have not yet been resolved, they have converged on the grammaticality of the phenomenon and the optionality of overt C2. Findings from experimental and psycholinguistic research, though few in number, have afforded the field a deeper understanding of the *why* and *when* behind C2 lexicalization. Only a few studies have investigated recomplementation structures: Casasanto and Sag (2008), Frank (2016), and Echeverría and López Seoane (2019). Additionally, Frank and Toribio (2017) will be reviewed in Chapter 3.

Casasanto and Sag (2008) predicted that the low complementizer *that* isn’t licensed by the grammar but lowers processing costs in complex constructions. In order to investigate this prediction, they investigated two fixed factors— length of the left-

dislocated material (one word versus seven words) and presence of the low *that* (null versus overt)— as represented in (2.6ab) below:

- (2.6) a. John reminded Mary that after he was finished with his meeting (that) his brother would be ready to leave.
- b. John reminded Mary that soon (that) his brother would be ready to leave.

Results from an acceptability judgment task and a self-paced reading task that measured the reading time of the critical region *brother* supported the idea of a tradeoff between grammaticality and processing complexity. Specifically, multiple *that* constructions are less acceptable but easier to process than their equivalent single *that* version. Notably, these judgments are conditioned by the length of the left-dislocated (LD) material, where the overt low *that* is more acceptable in the 7-word condition (e.g., 2.6a) than the 1-word condition (e.g., 2.6b). Furthermore, in the longer LD phrase condition, participants process the critical region *brother* faster in the overt low *that* condition. The authors argue that these results support a memory-based account of resolving processing difficulty (e.g., Gibson, 2000), where the low complementizer reiterates the information provided by the first and thus reduces the strain on working memory when it is spelled out. The overt complementizer further indicates that the left-dislocated segment has come to an end, which might also assist in the processing of an ensuing complement.

In the second study, Frank (2016) investigated the grammatical status of the low complementizer in Colombian Spanish. An aural acceptability judgment task, adopting the same scale as in the previous study, was designed to measure the acceptability of the overt low complementizer in question and statement contexts. Frank hypothesized that if Casasanto and Sag's (2008) findings apply to Spanish, then overt low complementizer constructions should be less acceptable than null ones. Furthermore, the author predicted that overt C2 question items would be rated higher than statement ones, given that C1 in

indirect questions is a reportative/quotative marker, which not only permits non-ask wonder verbs like *decir* ‘to say/tell’ to select for an indirect question but also is helpful in disambiguating a semantically ambiguous *wh*-complement (for a review see Cuza & Frank, 2011, 2015). For the reader’s convenience, question and statement test items from Frank (2016) are replicated in (2.7) and (2.8), respectively:

- (2.7) Me dijo que ese dibujo bonito que dónde iba a colgarlo.
‘S/he asked me where I was going to hang that beautiful painting.’
- (2.8) Me dijo que ese traje formal que iba a pedirlo.
‘S/he told me that s/he was going to order that formal suite.’

Results support the notion that the grammatical status of recomplementation is similar in English and Spanish. Specifically, the overt low complementizer in Spanish was indeed associated with a decrement in acceptability judgment, a result that was robust across statement and question constructions. Furthermore, no main effect was found for sentence type, which does not support the hypothesis that the type of dependency relationship between the complementizer and its complement influences acceptability ratings.

Finally, Echeverría and López Seoane (2019) created a corpus of 124 recomplementation instances in the 14th-century Spanish written text *El conde Lucanor*. Their analysis found that the length of the left-dislocated material was a significant predictor of C2 lexicalization, where C2 was more frequently spelled out when the intervening material between the C1 and C2 was longer (i.e., 1-3 words versus 4 or more words), see (2.9) and (2.10), respectively. This length effect has been documented in naturalistic present-day contexts as well. According to Radford (2018), the average number of words sandwiched between a high and low complementizer in his broadcast English corpus is 5.9.

(2.9) que algunos otros Ø non ayen envidia dellos
'that some others don't envy them'

(2.10) entendiendo que pues todo fincava en su poder, que podría obrar en ello como quisiese
'...understanding that because everything laid in his power, that he could act in it as he wished'

Lastly, Echeverría and López Seoane also found that mood, namely the subjunctive as opposed to the indicative, was a significant predictor of C2 lexicalization, adding credence to the hypothesis that the type of dependency relationship between the complementizer and its complement is a relevant factor in predicting C2 lexicalization (Frank, 2016). The authors conclude that conventional patterns of C2 usage can be predicted along probabilistic constraints rather than categorical rules.

In summary, the experimental evidence on recomplementation offers a more complicated story of C2 lexicalization than the optionality proposed by syntactic-theoretical accounts. Casasanto and Sag (2008) argue that C2 lexicalization is associated with a grammatical violation that is overridden by the benefit it brings to real-time sentence processing. Frank (2016) does not go as far as to claim that overt C2 is ungrammatical but does provide evidence that it is associated with a decrement in acceptability when compared to the null counterpart. Taken together, these two studies suggest that divergent outcomes in Spanish-English bilingual performance cannot likely be accounted for by turning to direct cross-linguistic influence effects. Nor does it seem that a decrement in acceptability judgment is motivated by prescriptive reflection, given recomplementation is not taught in the classroom. Finally, Echeverría and López Seoane (2019) adopt a usage-based account and ignore the question of grammaticality altogether. They note that C2 lexicalization occurs in written contexts and its occurrence can be predicted probabilistically along constraints like length of dislocated material and mood.

These three studies provide the field with a deeper understanding of why C2 is lexicalized in some contexts and not in others. It remains to be seen whether these C2 lexicalization patterns hold in heritage speaker populations or whether the extra burden of holding two languages in parallel while communicating in one's less dominant language triggers divergent outcomes. The following section will review some of the sources of divergence in heritage grammar and conclude by framing the acquisition of recomplementation within the model of divergent attainment.

2.3 SOURCES OF DIVERGENT ATTAINMENT

2.3.1 Vulnerable Phenomena

Heritage speakers of Spanish have been shown to diverge from the monolingual norm across several grammatical properties, including but not limited to subject-verb inversion (e.g., Cuza, 2016), number and gender agreement (e.g., Cuza & Pérez-Tatam, 2016; Scontras, Polinsky & Fuchs, 2018), *pro*-drop (e.g., Montrul, 2002, 2004, 2008, 2016), mood selection (e.g., Giancaspro, 2017; Perez-Cortes, 2016), clitic expression (e.g., Cuza et al., 2013; Montrul, 2010), as well as higher structural projections in the C-domain more generally (e.g., Bruhn de Garavito, 2002; Cuza, 2013; Cuza & Frank, 2011, 2015; Montrul, 2010; Silva-Corvalán, 1993). C-domain or left periphery phenomena like recomplementation or verb-second, embedded clauses and *wh*-questions are arguably more prone to divergence due to the complexity associated with the interface between syntax and pragmatics (e.g., Sorace, 2000, 2005, 2011).

Only a few studies have investigated the acquisition of the complementizer *que* 'that' in US heritage Spanish specifically. For example, on the topic of argument clauses,

Silva-Corvalán (1993) found several examples of null *que* in Los Angeles heritage speakers, as in (2.11):

- (2.11) Yo creo Ø inventaron el nombre.
'I think (that) they invented the name.'

Because the null complementizer is perfectly acceptable in the English equivalent, cross-linguistic influence (CLI) effects in the direction of the minority language may be in play.¹ Critically, no examples of *que* omission in relative clauses were found in the dataset, where omission is ungrammatical in Spanish but grammatical in English. Thus, the minority language appears to be susceptible to the effects of CLI when the surface structure of the two languages overlaps (e.g., Müller and Hulk, 2001; Yip and Matthews, 2009). This accounts for the observed *que* omission in Spanish argument clauses and lack thereof in relative clauses.

Cuza and Frank (2011, 2015) ask whether late second language learners and heritage speakers of Spanish of comparable high proficiency acquire the features that regulate the representation of simple indirect constructions, in which the overt complementizer is argued to be obligatory:

- (2.12) Me dijo (que)* cuándo iban a salir.
'He asked when they were going to leave.'

Data collected from an elicited production task, an acceptability judgment task, and a forced-choice preference task suggest that both bilingual groups produce and accept the null *que* condition in contexts that require a question interpretation. However, heritage speakers outperform second language learners, as demonstrated by greater overt complementizer production, higher acceptability rating of overt complementizer items, and a preference for the overt item. For example, when forced to choose between an overt and

¹ However, the narrative is more complicated since *que* omission can be found in similar Spanish contexts, such as *Te ruego (que) me lo envíes pronto* "I beg (that) you send it to me promptly".

a null complementizer option, second language learners overwhelmingly prefer the null condition, while 10/17 heritage speakers prefer the overt complementizer condition. While there is a certain level of individual variation within the heritage speaker group, when compared to the baseline group, heritage speakers overall produce, accept, and prefer null *que* constructions at a higher rate.

In summary, speakers of Spanish as a heritage language diverge from the monolingual norm across several grammatical properties. The few studies that have investigated the acquisition of the complementizer *que* ‘that’ in US heritage Spanish support the hypothesis that left periphery phenomena are a vulnerable domain. The following section introduces the Model of Divergent Attainment in Heritage Grammar (Polinsky & Scontras, 2020) and situates the acquisition of recomplementation within said model.

2.3.2 Model of Divergent Attainment

As discussed in the introduction to this chapter, Polinsky and Scontras (2020) lay out a preliminary framework for modeling the divergent morphosyntactic properties of heritage languages. With the proposed model, their goal is to accurately predict divergent outcomes in heritage language competence. They organize existing literature on divergent attainment along four intersecting categories of ‘problems’ or sources of divergence and propose specific triggers and outcomes of divergence.

The first source of divergence—the morphology problem—pertains to number and gender agreement and overmarking (e.g., past tense *sorteded* instead of *sorted*) or overregularization (e.g., past tense *bringed* instead of *brought*) (e.g., Polinsky, 2018; Scontras et al., 2018). The distance problem speaks to the challenges associated with long

distance dependencies (e.g., antecedent-gap, anaphor binding, agreement, left-dislocation) (e.g., Kim et al., 2010; Polinsky, 2011). The general outcome is a preference for local over non-local dependency even when this results in non-target performance. The third type of divergence is manifested as the silent problem, which refers to the challenges associated with the interpretation of null elements. For example, while Spanish is a *pro*-drop language, heritage speakers have been shown to prefer and overuse overt pronouns when compared to a relevant baseline group (e.g., de Prada Pérez, 2009; Montrul, 2016; Silva-Corvalán, 1994). Further, as discussed earlier, Cuza and Frank (2011, 2015) found that heritage speakers produce and accept the nontarget-like null *que* condition in contexts that require a question interpretation (see example 12). This suggests that silent material can be the source of reanalysis or restructuring of interpretive possibilities. Finally, the fourth source references the complexities associated with one-to-many mapping between form and meaning. Take for example scope ambiguity. Scontras et al. (2017) found that English-dominant heritage speakers of Chinese only allow surface interpretations of doubly-quantified sentences like *A shark attacked every pirate* in Chinese (target performance) and English (nontarget performance). These four problems help to focus the future research program that will follow the model.

Polinsky and Scontras (2020) propose the following triggers for divergent attainment and initial divergent outcomes, each considered in turn.

Triggers for divergent attainment

- (i) Quantity and quality of input
- (ii) Demands on processing and memory

Proposed divergent outcomes

- (i) Avoidance of ambiguity
- (ii) Resistance to irregularity
- (ii) Shrinking of structure

Much has been said on the differing experience between heritage speakers and relevant baseline groups (e.g., see Unsworth, 2016 for a review). A typical Spanish heritage speaker growing up in the United States may acquire Spanish as their first language in the home. But at school age, they enter an education system and society where English is the dominant language. Their quantity of input in Spanish is greatly reduced at this point. Over time, with an increased use of and exposure to English in school, social and work settings, the minority language becomes less dominant than the majority one. This experience stands in stark contrast to that of monolingual Spanish speakers. What's more, heritage speakers have fewer speaking partners in Spanish (e.g., immediate family, extended family, neighbors) as compared to their majority language partners. Importantly, this is not to say that the input from Spanish varieties that exhibit contact-induced changes or signs of attrition is less legitimate (e.g., Pascual y Cabo & Rothman, 2012). Rather, as Polinsky and Scontras (2019) note, "...increased exposure to the heritage language will only get heritage speakers so far; they also need exposure from a variety of sources" (11). Thus, one trigger for divergent attainment is the interrelated dimension of quantity and quality of input.

Processing pressure presents a second trigger for divergent attainment. We know from psycholinguistic research on monolingual populations that our online processing resources are limited and some areas of language comprehension and production test these limitations more than others. Examples include but are not limited to dependencies at a distance, the recovery of missing information, surprisal theory and the time/strength hypothesis (e.g., Arnold et al., 2000; Bailey & Ferreira, 2003; Gibson, 1998; Grodner & Gibson, 2005; Hale, 2001; Levy, 2008). These areas of high cognitive demand should be particularly difficult for the heritage speaker, who must maintain two grammars in parallel and communicates in their less dominant language (e.g., Montrul, 2016; Keating et al., 2016, Polinsky & Scontras, 2020; Sánchez, 2019). Sánchez (2019) proposes a bilingual

alignments hypothesis, where co-activation of stored information from different language components is particularly costly. These so-called ‘permeable’ alignments are possible across all levels of proficiency, though more likely at lower levels.

The Model of Divergent Attainment also offers three initial predictions of divergent outcomes, which are meant to be broad in nature: avoidance of ambiguity, resistance to irregularity and shrinking of structure. It’s unlikely that these three categories will sufficiently account for all present and future findings. However, many of the findings highlighted in the literature on divergent attainment reviewed above (e.g., one-to-one form to meaning mappings, overregularization, avoiding or reanalyzing of null elements, privileging shorter dependencies), are predicted by these three broad outcomes. As the authors point out, the specifics of the model and its predictions are a work in progress.

The phenomenon of recomplementation is a perfect candidate for the continued development of this model. First, unlike core aspects such as agreement, the phenomenon is not reinforced in school. Secondly, potential divergent behavior in Spanish-English bilinguals cannot be fully accounted for by cross-linguistic influence effects. As mentioned earlier, previous research argues that both Spanish and English monolinguals find the null complementizer more acceptable (Casasanto & Sag, 2008; Frank, 2016) and a single syntactic-theoretical analysis can account for the phenomenon cross-linguistically (Villa-García, 2019). Finally, it fits neatly into the model’s constrained research program. Recomplementation exemplifies the intersection of the silent problem and the distance problem. Specifically, C2 can be silent, requiring the interpretation of a null element. When C2 is spelled out, it serves as a reminder of the force of the embedded clause (Demonte & Fernández-Soriano, 2014; Martín-González, 2002). In other words, overt C2 functions as a secondary subordinator triggered by processing complexity. Further, C1 (i.e., the primary subordinator) is separated from C2 by the dislocated material. Retrieval of this

subordination information in the case that C2 is silent increases demands on working memory and processing resources (Casasanto & Sag, 2008).

In summary, we propose a multiple representations account of recomplementation. Specifically, while the TopicP account is the primary syntactic-theoretical account adopted in the literature for monolingual Spanish, we propose that the alternative DoubledForceP account better represents heritage speaker grammars. This divergent grammar is motivated by the linguistic complexity of recomplementation associated with silent elements and distance dependency. This, along with processing burden, leads to divergent attainment over time. With this framing of the acquisition of recomplementation in mind, the following section introduces the specific research questions that drive the remainder of the study.

2.3.3 Research Questions

The present study is an initial investigation of the acquisition of recomplementation structures in a heritage language population. It offers a contribution to the growing literature arguing that the left periphery is a vulnerable domain and elaborates on the Model of Divergent Attainment in Heritage Grammar (Polinsky & Scontras, 2020). We propose the following research questions:

- (RQ1) Do advanced heritage speakers accept the null C2 construction at a higher rate than the overt C2 option? Does language use or proficiency predict this outcome?
- (RQ2) Do advanced heritage speakers prefer the null C2 construction at a higher rate than the overt C2 option? Does language use or proficiency predict this outcome?
- (RQ3) With respect to (RQ1) and (RQ2), do advanced heritage speakers diverge from the monolingual baseline group?

With respect to (RQ1) and (RQ2), we predict that advanced speakers of Spanish as a heritage language will accept and prefer the overt C2 construction at a higher rate when compared to the null variety. Replenishment as the intersection of silence and distance problems combined with the extra burden of holding two languages in parallel while communicating in one's less dominant language will drive this effect (e.g., Sánchez, 2019). Specifically, according to the DoubledForceP account, a lexicalized C2 reintroduces the force or the semantic function of the complement. When C2 is not spelled out, the relevant semantic information must be interpreted or retrieved, which increases the burden on processing resources. We further predict that language use and proficiency will be correlated with rate of acceptability and proportion of preference of the overt C2 option. Specifically, heritage speakers with higher rates of language use and higher levels of proficiency will have more available resources for storage and retrieval of information and thus favor the overt variety.

With regard to (RQ3), we anticipate that the test group's performance across the acceptability judgment and preference tasks will not pattern the control group's behavior. Specifically, the null C2 variety will be associated with a decrement in acceptability judgment in the former group when compared to the latter. Further, the heritage speaker group will prefer the overt C2 option at a significantly higher rate when compared to the baseline group. Our first point of evidence comes from previous research which has shown that the overt C2 construction is associated with a decrement in acceptability judgment in a monolingual Spanish baseline group (Frank, 2016). Secondly, while a lexicalized C2 does under certain conditions bring a benefit to the relevant baseline group in real-time sentence processing, the conditions are not met for C2 lexicalization (e.g., long intervener length). Further, the offline measures are meant to serve as a window into language competence not sentence processing. The predicted differential outcomes for the two

groups support earlier research that has attested to the vulnerability of CP-related phenomena in bilingual populations. The prediction also supports a multiple representations account of recomplementation, where the TopicP and the DoubledForceP accounts pertain to the monolingual and bilingual groups, respectively. That is to say, the silence and distance problems along with the extra burden on processing leads to a reanalysis of the linguistic phenomenon and eventual divergent attainment (i.e., different representations) in heritage grammar.

2.4 THE STUDY

2.4.1 Participants

In order to test these predictions, a total of 27 participants took part in the present study. The participants were divided into two groups, a baseline or control group of native speakers of Colombian Spanish ($n=12$) and a US heritage Spanish test group ($n=15$). The former group was recruited through word of mouth with the support of local contacts in Bogotá, Colombia. They were all residents of Bogotá at the time of testing and had never lived in a country where a language other than Spanish was the primary language of society and education. Their ages ranged from 18-35 ($M= 22$, $SD= 5.4$) and they were all at minimum high school educated, with the majority having attended college (8/12). The majority of the control group were students, while other professional industries included engineering, logistics and music. The test group was recruited from a large public university in the southwestern United States. The participants' birthplace was the United States, with one exception (Mexico). They were all raised by native Spanish speaking parents who spoke to them in either Spanish (8/15) or Spanish and English (7/15) in childhood. Their primary language of instruction from primary school through college was

English. Lastly, their ages ranged from 18-23 ($M= 20$, $SD=1.5$). Given that recomplementation is not thought to be a source of dialectal variation, we adopted a control group that does not directly match the input of the heritage speaker group. However, this incongruency along with the possible inflation of type II error owing to a low number of participants constitute two of the study's limitations.

The last portion of the language history questionnaire (see Appendix A) included a brief proficiency self-assessment in both Spanish and English, which facilitated a language dominance calculation. Ratings along a four-point scale (i.e., 1 “basic”, 2 “adequate”, 3 “good”, 4 “excellent”) were elicited across four modes: reading, writing, speaking and comprehension. As reported in Table 1, after collapsing the four modes, the test group's mean for English is greater than Spanish, $M= 3.83$ and $M= 2.97$, respectively. Importantly, it is also true, with one exception, that each participant was English dominant. The one exceptional participant self-rated as balanced. Contrastingly, the baseline's mean for Spanish is greater than English ($M= 3.83$ and $M= 2.97$, respectively) and without exception, each participant was Spanish dominant. For a summary of the profile of each group see Table 2.1.

Table 2.1. Participant profile

Selected Information	Baseline Group	Test Group
Birthplace	Colombia	USA (1 MEX)
Gender	7 male, 5 female	9 male, 6 female
Age at testing	<i>M</i> = 22, <i>SD</i> = 5.4	<i>M</i> = 20, <i>SD</i> =1.5
Level of Education	8 college, 4 high school	College
Language Spoken as a Child	SPAN	8 SPAN, 7 Both
Primary Language of Instruction	SPAN	ENG
SPAN Proficiency (DELE out of 50)	<i>M</i> = 43, <i>SD</i> = 2.7	<i>M</i> = 37, <i>SD</i> = 4.9
SPAN/ENG Self-Reported Proficiency (1-basic to 4-excellent)	SPAN: <i>M</i> = 3.79, <i>SD</i> = 0.35 ENG: <i>M</i> = 1.79, <i>SD</i> = 0.60	SPAN: <i>M</i> = 2.97, <i>SD</i> = 0.27 ENG: <i>M</i> = 3.83, <i>SD</i> = 0.28
Dominant Language	SPAN	ENG (1 balanced)
More Comfortable Language	SPAN	9 ENG, 6 Both

In addition to the language history questionnaire, all participants completed an adapted version of the DELE (Diploma of Spanish as a Foreign Language) proficiency test, as in Appendix B (e.g., Montrul & Slabakova, 2003). As demonstrated in Table 2.1 above, the control group scores were in the range of 39-47 (*M*= 43, *SD*= 2.7) and the test group was in the range of 30-44 (*M*= 37, *SD*= 4.9) out of a possible 50 items. An independent-sample *t*-test showed a significant difference between the baseline and test group proficiency results, ($t(24) = -1.882, p < .001$). Specifically, the native speakers of Colombian Spanish outperformed the US heritage Spanish test group.

Lastly, for the test group, a proportion of current Spanish language use was calculated. Participants were asked how often they used Spanish in the following four contexts: school, home, work, and social situations. Their responses were recorded along

the scale of 0 “English only”, 1 “mainly English”, 2 “a little more English”, 3 “both equally”, 4 “a little more Spanish”, 5 “mainly Spanish”, 6 “Spanish only”. Individual responses were then divided by 6 (Spanish only) to calculate a proportion of Spanish use, with a possible range of 0 “English only” to 1 “Spanish only”. As demonstrated in Table 2.2, the overall proportion of Spanish language use is $M= 0.24$ ($SD= 0.13$) or one quarter of total language use.

Table 2.2. Test group proportion Spanish language use

Spanish Use	Proportion	Interpretation
at school	$M= 0.10, SD= 0.11$	Mainly ENG
at home	$M= 0.47, SD= 0.32$	SPAN & ENG equally
at work	$M= 0.18, SD= 0.18$	Mainly ENG
in social situations	$M= 0.23, SD= 0.15$	Mainly ENG
overall	$M= 0.24, SD= 0.13$	Mainly ENG

The imbalance between Spanish and English is particularly driven by school, work and social contexts. Interestingly, the proportion of Spanish used in the home is $M= 0.47$ ($SD= 0.32$), suggesting nearly equal usage of Spanish and English. This is perhaps not surprising given the students are college-aged and “home” still refers to their family home, where they were raised with either Spanish or Spanish and English.

2.4.2 Methods and Design

The study itself consisted of two experiments, an aural acceptability judgment task (AJT) and a written forced-choice preference task. Supplementing reading/writing tests with aural/oral ones, or avoiding the written medium altogether, has been argued to be

critical when eliciting data from heritage language communities. Written tasks underrepresent the overall performance abilities of heritage speakers (e.g., Bowles, 2011; Cuza & Frank, 2015; Cuza, 2013; Potowski, Jegerski, & Morgan-Short, 2009). For example, within the same study, heritage speakers have been outperformed by second language learners in written tasks and then outperformed their counterparts in verbal ones (e.g., Alarcón, 2011; Montrul, 2011; Montrul, Foote, & Perpiñán, 2008).

The aural AJT was designed on a professional package of the *Weebly* web-hosting service. All testing instructions and tokens were read by a naïve native Spanish speaker, whose voice was recorded and then edited on version 2.0.3 of Audacity® software (<https://audacityteam.org/>). The recorded instructions explained that each test item contained three sections: a preamble, a question, and a response to the question. After listening to all three parts, the listener's task was to determine whether the response was well formed, using the scale 1-totally acceptable to 7-totally unacceptable, adopting Casasanto and Sag's (2008) scale for maximal comparison. Two examples of the aural test items are represented and translated below:

(2.13) *Question Condition*

- Preamble: Ese dibujo bonito, ¿dónde vas a colgarlo?
'Where are you going to hang that beautiful picture?'
- Question: ¿Qué te dijo Susana?
'What did Susana ask you?'
- Response: Me dijo que ese dibujo bonito (que) dónde iba a colgarlo.
'She asked me where I was going to hang that beautiful picture.'

(2.14) *Statement Condition*

- Preamble: Ese traje formal, voy a pedirlo.
'I am going to order that formal suite.'
- Question: ¿Qué te dijo Susana?
'What did Susana tell you?'

Response: Me dijo que ese traje formal (que) iba a pedirlo.
'She told me that she was going to order that formal suite.'

As demonstrated in (2.13) and (2.14), the preamble represents a clitic left-dislocated (CLLD) statement and question, respectively, establishing the argument (e.g., *el traje, el dibujo*) as having been previously mentioned in the discourse. The left-dislocated segment is specifically composed of a demonstrative adjective, a noun, and an adjective that modifies the noun, and is followed by an informal future expression with the clitic attached to the infinitive verb. The preamble is followed by the question, which remains consistent across all test items. The third and final section pertains to the response, which is the recomplementation test item that participants are to judge on the aforementioned ordinal scale. Test items include embedded statements and questions with a variety of sandwiched arguments previously mentioned in the discourse. The material that intervenes between C1 and optional C2 is controlled at three words. Finally, for the question condition only, the locative or the temporal adjunct *wh*-words *dónde* 'where' and *cuándo* 'when', respectively, are inserted before the *wh*-complement.

The aural AJT was designed to measure the acceptability of the null versus the overt low complementizer in both question and statement contexts. As demonstrated in Appendix C, it is composed of 6 statements with null secondary complementizer items, 6 statements with overt secondary complementizer items, 6 questions with null secondary complementizer items, and 6 questions with overt secondary complementizer items for a total of (n=24) tokens. These test tokens were scrambled with 48 distractor items (of both question and statement varieties investigating unrelated phenomena) so as to ensure that no two identical conditions appeared consecutively.

In the second experiment, the paper and pencil forced-choice preference task, participants were directed to read a short preamble and then select the preferred one of two available continuation statements, see (2.15) and (2.16) below:

(2.15) *Question Condition*

Ayer Miguel tuvo que recordarme de la chaqueta que habíamos visto.

‘Yesterday, like last week, Miguel reminded me about the jacket that we had seen.’

___ Miguel me repitió que esa chaqueta, que cuándo iba a comprarla.

___ Miguel me repitió que esa chaqueta, cuándo iba a comprarla.

‘Miguel asked me again when I was going to buy that jacket.’

(2.16) *Statement Condition*

Ayer Leonardo tuvo que recordarme del folleto que creamos la semana pasada.

‘Yesterday Leo reminded me of the flyer that we created last week.’

___ Leonardo me dijo que ese folleto, iba a distribuirlo en el centro.

___ Leonardo me dijo que ese folleto, que iba a distribuirlo en el centro.

‘Leo told me that he was going to distribute that flyer downtown.’

As in the AJT, recomplementation test items include embedded question and statement sentence types for a total of 16 tokens. Test items were composed of a variety of left-dislocated topics that are previously mentioned in the discourse and the material that intervenes between C1 and optional C2 is controlled at two words: a demonstrative adjective and a noun (see Appendix D). Test items were scrambled with 24 distractors and the two choices within each token were then counterbalanced in order to avoid a bias for selecting the first available option. Given the forced-choice design of the experiment, the dependent measure was binary, where a preference for the null option was coded with a score of 0 and a preference for the overt option received a score of 1.

2.4.3 Results

The baseline and test group results from the aural AJT are displayed in Figures 2.1 and 2.2, respectively. Each column represents the mean acceptability rating with standard error bar for each of the four test conditions: null C2 questions, null C2 statements, overt C2 questions and overt C2 statements. Recall that each condition was made up of six test items for a total of (n=24) tokens. Further, the acceptability rating scale is a seven-point scale from 1-totally acceptability to 7-totally unacceptable. That is to say, the shorter the column, the more acceptable the condition. The scale has been truncated on the y-axis to better fit the data.

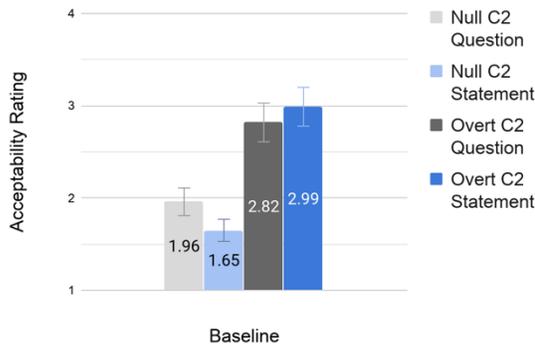


Figure 2.1. Baseline group mean acceptability ratings

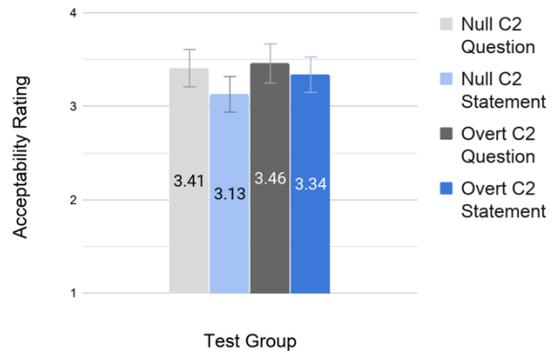


Figure 2.2. Test group mean acceptability ratings

The baseline group results from Figure 2.1 suggest that the null C2 variety, independent of sentence type, is judged as more acceptable than the overt C2 variety. This behavior appears to contrast with the behavior of the test group. Specifically, Figure 2.2 shows similar behavior across all conditions for the test group. Table 2.3 summarizes these results.

Table 2.3. Mean acceptability ratings from 1 “totally acceptable” to 7 “totally unacceptable”

	Null C2 Question	Null C2 Statement	Overt C2 Question	Overt C2 Statement
Baseline group	1.96 (<i>SE</i> = 0.15)	1.65 (<i>SE</i> = 0.12)	2.82 (<i>SE</i> = 0.21)	2.99 (<i>SE</i> = 0.21)
Test group	3.41 (<i>SE</i> = 0.20)	3.13 (<i>SE</i> = 0.19)	3.46 (<i>SE</i> = 0.21)	3.34 (<i>SE</i> = 0.19)

In order to investigate these descriptive statistics and to shed light on (RQ1) and (RQ3), a logistic mixed effects model for ordinal data was run with the CLMM (cumulative link mixed model) function in the R ordinal library (Christensen, 2015; R Core Team, 2017). The model defined three fixed effects-- group (Colombian or heritage speaker), type (question or statement) and C2 (null or overt)-- four interactions (group*type, group*comp, type*comp, group*type*comp) and one random intercept for subject. Both group ($\beta = 2.236, z = 3.331, p < 0.001$) and C2 ($\beta = 1.339, z = 4.045, p < 0.001$) were significant, along with the interaction of group*C2 ($\beta = -1.307, z = -3.044, p = 0.002$). Type did not reach significance ($\beta = -0.563, z = -1.626, p = 0.104$). To further explore the interaction, a post hoc pairwise comparison with Bonferroni adjustment was run. The analysis showed that the effect of C2 was significant in the Colombian baseline group ($\beta = -1.308, z = -5.939, p < 0.001$) and not the heritage speaker test group ($\beta = -0.093, z = -0.498, p = 0.619$). In summary, the results demonstrate that heritage speakers do not accept the null C2 construction at a higher rate than the overt C2 option, partially confirming what we predicted in (RQ1). The confirmation is only partial because rather than accepting the overt C2 construction at a higher rate, the effect of C2 was not significant. The results also demonstrate that the test group’s behavior diverges from the baseline group, who rates the

overt C2 variety with a decrement in acceptability, confirming our prediction of divergent performance in (RQ3).

In an attempt to further analyze the test group's divergent behavior and individual variation, three correlations were run with data from the participant profile (Tables 2.1 and 2.2). We investigated whether Spanish proficiency (results from DELE exam), degree of English dominance² or Spanish language use predicted the outcome of the null C2 acceptability ratings, where the divergence is most salient (see Figures 2.1 and 2.2). We hypothesized that Spanish proficiency, dominance and language use as proxies for experience might predict the test group's divergent behavior and be positively correlated with the acceptance of the null C2 variety. However, results returned weak and insignificant correlations between the acceptability ratings and proficiency ($r = 0.105, p = 0.708$), dominance ($r = -0.022, p = 0.937$) and language use ($r = -0.031, p = 0.914$).

A final grouping analysis was adopted to determine whether the descriptive and statistical analyses above were also observed at the individual level. In this analysis, participants were divided into five distinct groups, depending on (1) whether they rated the null or the overt variety as more acceptable and (2) the degree by which they did so. Specifically, those who favored the null over the overt variety by a difference of less than 1 on the 1-7 rating scale, were placed in the "null low" group. If the difference was greater than 1, then they were placed in the "null high" group. The same divisions were applied to those who favored the overt variety. Finally, if the difference was equal to zero, they were placed in the "equal" group. Sentence type (question or statement) was collapsed since the effect was not significant. The results of this analysis for Colombian (baseline) and US heritage (test) groups are represented in Table 2.4.

² A dominance coefficient was calculated by dividing the collapsed Spanish self-rating from the collapsed English self-rating for each participant. The potential range of variation of English dominance is 0.25 ("basic" in Spanish and "excellent" in English) to 1.00 (balanced).

Table 2.4. Participants who favored null or overt C2 by a low/high amount³

	Null		Overt		Equal
	low (<1)	high (>1)	low (<1)	high (>1)	
Baseline group	2/12 (17%)	6/12 (50%)	2/12 (17%)	0/12 (0%)	2/12 (17%)
Test group	7/15 (47%)	1/15 (7%)	6/15 (40%)	0/15 (0%)	1/15 (7%)

As demonstrated above, the majority of the baseline Colombian participants rate the null variety as more acceptable, 8/12 (67%). Furthermore, the majority of the Colombian participants who find the null varieties more acceptable fall in the “high” category. Curiously, this grouping analysis shows us that the heritage speaker participants also rate the null variety as more acceptable by a small majority, 8/15 (53%). However, Table 2.4 suggests a much more complicated story for the test group. The data appears to be distributed bimodally, with 8/15 (53%) participants accepting the null variety at a higher rate and 6/15 (40%) accepting the overt variety at a higher rate. Importantly, nearly all the test group participants are housed in the “low” category. This suggests that despite the fact that there is only one US heritage participant in the “equal” category, these participants do not find the overt or the null varieties much more acceptable than one another.

In order to supplement the AJT findings, participants also completed a forced-choice preference task. Recall that participants are instructed to select the preferred one of two available options (null and overt C2 varieties). Thus, the dependent measure was binary, where a preference for the null option was coded with a score of 0 and a preference for the overt option received a score of 1. Figure 2.3 below depicts the proportion of overt C2 options that were preferred, where any value below 0.50 entails null C2 preference.

³ Percentages in Tables 2.4 and 2.5 are rounded to the nearest whole number. This leads to some minor discrepancies in totals for some rows.

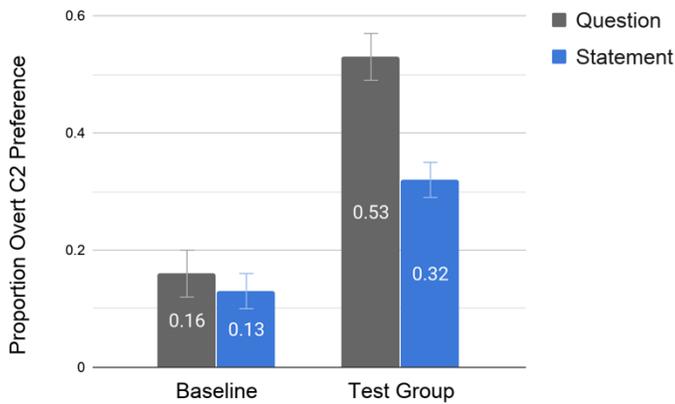


Figure 2.3. Proportion overt C2 preference

In Figure 2.3, the baseline group demonstrates a strong preference for null option that is robust across question and statement conditions, 0.16 and 0.13, respectively. Contrastingly, the test group not only demonstrates a marginal preference for the overt C2 in question items, but this behavior appears to differ with respect to a moderate preference for the null C2 in statement items. Table 2.5 summarizes these results.

Table 2.5. Proportion overt C2 preference

	Question	Statement
Baseline group	0.16 (<i>SE</i> = 0.04)	0.13 (<i>SE</i> = 0.03)
Test group	0.53 (<i>SE</i> = 0.05)	0.32 (<i>SE</i> = 0.04)

In order to elaborate on these preliminary findings and to shed light on (RQ2) and (RQ3), a general linear mixed effects model was run with the GLMER function in R (R Core Team, 2017). The model defined two fixed effects, group (Colombian or heritage speaker) and type (question or statement), an interaction for group*type and a random intercept for subject. Group was found to be significant ($\beta = 2.714$, $z = 2.833$, $p = 0.004$),

while type was not ($\beta = -0.328, z = -0.699, p = 0.485$). The interaction of the two was marginally significant ($\beta = -1.156, z = -1.934, p = 0.053$). In sum, the significant effect of group tells us that the heritage speaker group's proportion of overt C2 preference is significantly greater than the baseline group's, suggesting divergent behavior and confirming our expectations for (RQ3). Preliminary results do however point to the test group's marginal overall preference for the null variety when both sentence types are averaged together (0.43). Thus (RQ2), which asks whether heritage speakers prefer the null variety over the overt variety deserves further scrutiny.

To explore the marginal interaction between group and type, we ran a post hoc pairwise comparison for all combinations of group and type with Bonferroni adjustment. This analysis shows that the effect of type is significant in the test group ($\beta = 1.484, z = 4.001, p < 0.001$) and not in the baseline group ($\beta = 0.328, z = 0.699, p = 0.485$). This curious effect can be explained if we consider the complementizer in embedded questions as being helpful in disambiguating a semantically ambiguous *wh*-complement (for a review see Cuza & Frank, 2011, 2015). The problem of distance introduced by left-dislocated material intervening between C1 and optional C2 can make the lexicalization of C2 all the more helpful.

As a continuation of the post hoc analysis, two moderate and statistically significant correlations were found with language use as a predictor, while the remaining correlations with the proficiency and dominance predictors were weak and insignificant. Specifically, the test group's proportion of Spanish language use predicted both the proportion of overt C2 preference in questions ($r = 0.525, p = 0.045$) and the overt C2 preference overall ($r = 0.595, p = 0.019$). This means that heritage speaker participants who use Spanish more often in the school, work, home and social contexts prefer the overt C2 items at a higher rate in question and combined question and statement conditions. *Prima facie* it is

surprising that Spanish language use is positively correlated with behavior that is less baseline-like. However, this finding is entirely compatible with a divergent attainment and multiple representations account, where heritage speaker grammars differ from baseline grammars across certain domains for reasons that are not exclusively related to experience.

One last grouping analysis was designed to determine whether the descriptive and statistical analyses above were also observed at the individual level. Specifically, participants were grouped into three categories: those who preferred the null C2 variety (proportion of overt C2 preference between 0.0 and 0.49), those who preferred the overt C2 variety (proportion of overt C2 preference between 0.51 and 1.00) and those who displayed no preference (proportion of overt C2 preference equals 0.50). Type (question or statement) was not collapsed as it was for the AJT, given the effect proved significant in the preference task.

Table 2.6. Participants who preferred null or overt C2 overall

	Type	Null	Overt	Equal
Baseline group	Question	10/12 (83%)	1/12 (8%)	1/12 (8%)
	Statement	12/12 (100%)	0/12 (0%)	0/12 (0%)
Test group	Question	5/15 (33%)	7/15 (47%)	3/15 (20%)
	Statement	11/15 (73%)	4/15 (27%)	0/15 (0%)

As demonstrated by Table 2.6, the baseline participants prefer the null C2 variety over the overt one. This result is robust across both question (10/12, 83%) and statement types (12/12, 100%). Contrastingly, individual variation is high in the test group. The heritage speaker group roughly patterns the baseline group in the statement condition, with the majority of participants demonstrating null C2 preference (11/15, 73% vs. 4/15, 27%).

Still, it is important to note that nearly 1 out of 3 participants do prefer the overt C2 statement condition. In the question condition, the test group participants vary considerably, 7/15 (47%) prefer the overt option, 5/15 (33%) prefer the null option and 3/15 (20%) have no preference. Where the AJT test group analysis was described as bimodal, for the preference task, high individual variation appears to be the most accurate conclusion.

In sum, the heritage speaker group prefers overt C2 at a significantly higher rate than the baseline group. Further, their proportion of overt C2 preference is significantly greater in questions as compared to statements. Importantly, an individual grouping analysis complicates this narrative by showing high individual variation. Curiously, the effect of type (statement vs. question) found in the preference task was not replicated in the AJT. The preference task was seemingly more sensitive to this divergent outcome.

2.5 DISCUSSION

The present study offered an initial exploration of recomplementation in advanced speakers of Spanish as a heritage language. We adopted an aural version of an acceptability judgment task so as not to underrepresent the overall performance abilities of this population. This task was supplemented by a forced-choice preference task. Together these offline tasks were assumed to serve as a window into heritage speaker knowledge of the secondary complementizer in statement and question contexts. Data analyzed at the group and individual level shed light on whether heritage speakers (1) accept and (2) prefer the null C2 variety at a higher rate than the overt C2 one and (3) whether their behavior diverged from a relevant baseline group. We predicted that the test group would accept and prefer the overt C2 at a higher rate and that this would diverge from the control group.

These predictions were primarily motivated by previous research on the vulnerability of CP-related phenomena in bilingual populations, sources of divergence that pertain to null elements and distance dependencies, and evidence of C2 lexicalization constraints in the relevant baseline group. The results are not categorical and paint a complex picture.

Findings from the AJT show that heritage speakers do not accept the null C2 construction at a higher rate than the overt C2 option. *Prima facie*, this supports our expectations. However, heritage speakers did not significantly accept the overt C2 at a higher rate either, though we cannot discount the possibility of type II error. Rather, there was no significant effect for C2. This result is supported by the individual analysis, where what appears to be a bimodal distribution at first glance-- 8/15 rate the null variety higher vs. 6/15 the overt variety-- turns out to be a bit misleading. That is to say, the amount that tips the scale for nearly all participants is “low”. Thus, we can conclude that the heritage group does not find the overt or the null varieties much more acceptable than one another. Importantly, no significant effect of language use or proficiency was found. Still, in line with our expectations, this behavior does diverge from the baseline group, which accepts the null variety at a significantly higher rate (see Figures 2.1 and 2.2). Individual analysis supports this finding, where despite some variation, the majority of the baseline participants rate the null variety as more acceptable.

The picture is made still more complex when we consider the preference task results. A group analysis of the heritage speakers displays a marginal overall preference for the null variety, 0.43, where a score below 0.50 is akin to null preference. In addition, there’s a significant effect for sentence type, where heritage speakers prefer the null variety significantly more in embedded statements than in embedded questions. In fact, we see a moderate preference for the null variety in statement condition (0.32) and a marginal overt variety preference in the question condition (0.53). Further, we find that Spanish language

use is positively correlated with overt C2 preference. This finding suggests that heritage speakers who use Spanish more often diverge most from the baseline group. Finally, results from heritage speaker individual analysis demonstrate high individual variation, particularly in the question condition, where 7/15 (47%) prefer the overt option, 5/15 (33%) prefer the null option and 3/15 (20%) display no preference. All this to say that, as in the AJT, the test group results from the preference task do not perfectly align with categorical expectations. Also, like the AJT, our expectations are confirmed with regard to divergent behavior when compared to the baseline group. The control group prefers the null variety at a significantly higher rate, an effect that is both robust across sentence type and confirmed by the individual analysis.

With respect to the overall marginal to no effect of C2 displayed by the heritage speaker group, we offer the following interpretation. The oversuppliance of overt forms (e.g., pronouns) among heritage speakers and bilingual populations more generally is well documented. Polinsky and Scontras (2020) devote a section to the silent problem with reference to *pro*-drop specifically. They note that the increase in the adoption of overt forms can be traced to earlier generations, even first-generation immigrants (e.g., Montrul, 2016; Otheguy & Zentella 2012; Otheguy et al., 2007; Sorace, 2004). What's more, the overuse of overt material has been observed as a result of contact itself, not fully explained by cross-linguistic influence effects (e.g., de Prada Pérez, 2009). Importantly, the claim is not that null forms have not been acquired, rather, their rates of use are reduced proportionally. In summary, there's no reason to assume that the probabilistic constraints for C2 lexicalization are equivalent across heritage speaker and baseline populations. It is possible that the constraints or features regulating the secondary *que* result in optional selection between null and overt C2 as the correct setting in the specific contexts of the test

items. This is not to say that specified constraints aren't evident in untested environments (e.g., at distance, in other dependencies).

While it has been argued that heritage speaker acquisition of C-domain phenomena (e.g., verb-second, embedded clauses and *wh*-questions) are more likely to diverge from the baseline group due to the complexity associated with the interface between syntax and pragmatics (e.g., Sorace, 2000, 2005), we would like to consider a different framework, namely the Model of Divergent Attainment in Heritage Grammar (Polinsky & Scontras, 2020). We argue that the silence and distance problems along with the extra burden on processing leads to a reanalysis of the linguistic phenomenon and eventual divergent attainment. According to the DoubledForceP account, a lexicalized C2 reintroduces the semantic function of the embedded clause. When C2 is null, it must be interpreted and/or relevant information must be retrieved from the primary complementizer (C1), which is separated by a string of intervening material seemingly limitless in length. As we know from monolingual processing literature, interpretation of missing elements and distant information retrieval places an increased demand on working memory and processing resources (e.g., Arnold et al., 2000; Bailey & Ferreira, 2003; Gibson, 1998; Grodner & Gibson, 2005; Hale, 2001; Levy, 2008; Casasanto & Sag, 2008). The strain on available resources should be even greater in bilinguals and heritage speakers specifically, who are holding multiple languages in parallel while working in the less dominant one (e.g., Montrul, 2016; Keating et al., 2016, Polinsky & Scontras, 2020). As a result, we propose a multiple representations account of recomplementation, where DoubledForceP is adopted by heritage speaker grammars and TopicP is adopted by monolingual grammars. In sum, we propose the economy of resources triggers the heuristics of avoid silent material and establish shortest distance dependency, which over time, leads to reanalysis and

restructuring across certain domains. This is a testable framework that accounts for the divergence in heritage speaker outcomes and informs heritage language acquisition theory.

We believe the pronounced effect of sentence type in the preference task adds credence to this interpretation. First of all, it is difficult to explain why the effect was not replicated in the AJT. We can only assume that a forced-choice situation, where both responses are presented before the participant, is more sensitive to this outcome. To be sure, the effect will need to be replicated in future studies. We speculate here as to why this effect may be realized. One such possibility is that the complement in embedded *wh*-questions, e.g., *dónde iba a colgarlo* ‘where I was going to hang it’ is semantically ambiguous [+QU] or [-QU] (see example 2.7; for a review see Suñer, 1993). This is not true of the embedded statements, e.g., *iba a pedirlo* ‘s/he was going to order it’, which are [-QU] (see example 2.8). We might then consider the complementizer in embedded *wh*-questions as being helpful in disambiguating a semantically ambiguous *wh*-complement, where the interrogative force and/or the reportative/quotative nature of the secondary complementizer predicts C2 lexicalization. Echeverría and López Seoane’s (2019) finding that mood predicts C2 lexicalization adds credence to the importance of dependency type. This problem of ambiguity, or one-to-many mapping between form and meaning in the *wh*-complement, may even be exacerbated by the distance between C1 and C2. This is an empirical question with distance as an independent factor.

In addition to providing an initial investigation on recomplementation in Spanish as a heritage language, this study aimed to contribute to the literature on the acquisition of the left periphery in bilingual populations, as well as provide new evidence towards a working model heritage language grammar. Earlier we argued that the heritage speaker group does not pattern with the baseline group and we supported this conclusion with robust evidence across tasks, significant effects and interactions therein. Importantly, as

our data has shown, the burden of demonstrating divergent attainment in recomplementation need not equate to behavior that is the opposite of the baseline group, such as a significantly higher acceptability rate or preference for the overt C2 when compared to the null variety. Rather, marginal preference either overt C2 or null C2 and the absence of a significant effect for C2 (question items in the preference task aside) is also a form of divergence. Thus, this finding offers new evidence in favor of the growing body of literature arguing that the left periphery is a “vulnerable” domain (e.g., Bruhn de Garavito, 2002; Cuza, 2013; Cuza & Frank, 2011, 2015; Frank, 2013; Montrul, 2010; Silva-Corvalán, 1993; Zapata et al., 2005). The next step for this research program is to investigate documented probabilistic constraints (e.g., distance) via online measures, while incorporating relevant socio- and psycho-linguistic variables (e.g., language use, input history, dominance, age of acquisition, working memory span) into the experimental design. One clear prediction is that participants with a low working memory span will benefit more from C2 lexicalization, particularly when the intervening material is long. Future studies should also incorporate production tasks in order to surface any potential mismatches between production and comprehension.

A broader contribution of this study is that it provides some clarity as to why the experimental investigation of heritage language populations is important for general theory. As we saw in Section 2.2.2, experimental methods can go beyond the question of C2 grammaticality and categorical rules of lexicalization, which do not adequately account for the phenomenon of recomplementation. Preliminary experimental evidence on monolingual populations has suggested overt C2 can be predicted probabilistically along defined constraints, such as length of dislocated material and mood (Casasanto & Sag, 2008; Echeverría & López Seoane, 2019). This claim would be strengthened with more corpus evidence. In the present study, we see that even probabilistic predictions along

defined constraints fail to fully capture the effect. Different populations, even individuals within the same population, can display probabilities and constraints that differ from one another (e.g., Dabrowska, 2012). Given already documented domains of divergent attainment and relatively high within group variation, heritage speaker populations afford researchers valuable data. And so, our general theory evolves from categorical rules, to probabilistic predictions along defined constraints, to probabilistic predictions along variable constraints that fluctuate based on experience and cognitive-related factors as an interrelated dimension. In Chapter 3, the experimental study of recomplementation is extended beyond bilingual US heritage speakers to understudied monolingual varieties.

Chapter 3: Cuban and Colombian Spanish Study

The previous chapter offered novel data on overt and null C2 patterns in US heritage Spanish. In so doing, it provided evidence that diverse speaker populations inform syntactic-theoretical accounts of recomplementation, as well as contribute to debates on the role of processing complexity in language development. Namely, we propose a multiple grammars account of recomplementation, where linguistic complexity associated with silent elements and distance dependency can lead to divergent attainment and changes in underlying representation. Importantly, US heritage Spanish as an understudied variety, also informs research on linguistic variation in the Spanish-speaking world. The present chapter extends the topics of understudied varieties and linguistic variation to the investigation of recomplementation as a potential locus of dialectal variation in Cuban and Colombian Spanish. Further, it expands on the previous chapter by adopting elicited imitation and sentence completion experimental methods, as well as a microvariational analysis, where recomplementation is investigated at the levels of intervener length, intervener type and sentence type.

3.1 INTRODUCTION

Generativist and Minimalist programs (Chomsky 1965 et seq.) have played a significant role in shaping our understanding of Grammar and how language is represented. Programs have historically relied on intuition and judgements of acceptability as the basis for advancing theoretical accounts. Importantly, the syntactic-theoretical literature has been biased towards Indo-European languages, such as English and Spanish, and selected varieties of the same. With reference to Spanish specifically, Peninsular Spanish has been held as the exemplar variety. One potential side effect of this is the tendency to apply

morphosyntactic descriptions of a few dialects to the many. This is problematic given the variation across the Spanish dialects of the world, from Peninsular to Latin American, from Caribbean to “Mainland” and the many microvariations therein. For example, by investigating morphosyntactic dialectal variation in lesser-studied Caribbean populations, we now know that traditional assumptions related to pronominal expression and free variation of subject-verb word orders had carried a Mainland bias (e.g., Camacho, 2013; Lipski, 1977; Martínez-Sanz, 2011; Orozco, 2015; Ortiz-López, 2009; Otheguy & Zentella, 2012; Toribio, 2000).

The linguistic phenomenon of recomplementation is no exception. In Villagarcía’s (2015) monograph *The Syntax of Multiple-que Sentences in Spanish*, he makes clear that his primary source of evidence is Peninsular Spanish. He states, “...I limit the discussion to Modern Castilian/European/Iberian/Peninsular Spanish,” and continues that “...although recomplementation may be more prevalent in certain dialects, I will not concern myself with addressing such potential dialectal differences here” (16). Furthermore, Fontana (1993) explains that recomplementation constructions with overt secondary *que* were the norm among the scribes of Old Spanish. This is most recently validated by Echeverría and López Seoane (2019) who found that 68 out of the 124 (55%) recomplementation structures in the 14th-century Spanish written text *El conde Lucanor* lexicalized secondary *que*. While recomplementation may be more restricted in modern Spanish, Fontana suggests that “...to this day, similar constructions are still possible in many Spanish dialects in a colloquial register” (163). For example, he points out that all native speakers that he has consulted accept (3.1).

- (3.1) Les prometieron **que** si votaban por ese candidato **que** les iban a dar mucho dinero.
them promised.3Pl that if voted.3Pl for that candidate that they were going.3Pl to give much money

‘They promised them that if they voted for that candidate, they would give them a lot of money.’

Demonte and Fernández-Soriano (2009) investigate five CP-related structures including recomplementation and argue that they are “...quite extended in all dialects of Spanish” (1). While the dialect of the consulted native speakers in these studies is not clearly disclosed (aside from the researchers’ own Peninsular Spanish), it is safe to surmise that the judges didn’t represent a random sample of the Spanishes of the world.

To counter a potential Peninsular bias in the literature, Frank and Toribio (2017) investigated recomplementation in Havana Cuban Spanish. The results from two experimental offline production tasks suggest that the secondary *que* neither facilitates complement integration nor is it licensed by the grammar. The authors’ conclusions remain speculative since no data was collected from speakers of other varieties for direct comparison, and only the variable of intervener length was analyzed. Thus, little can be said of the potential microvariations or social factors found therein. Still, crucially, although their preliminary finding is incongruent with the syntactic-theoretical literature, it is not unprecedented in the experimental literature. Casasanto & Sag (2008), a study on English in the US, argued for a grammaticality/processing tradeoff, where secondary *that* is associated with a grammatical violation that is overridden by the benefit it brings to sentence processing. Frank (2016), a study on Spanish in Colombia, highlights a decrement in acceptability in lexicalized secondary *que* when compared to the null variety. Together, these three studies point to a tension: On the one hand, syntactic-theoretical literature makes the case for the grammaticality/optionality of overt C2; on the other hand, experimental literature argues for its marginal acceptability or ungrammaticality. With this in mind, the specific aims of this chapter are as follows:

- (i) Address Frank and Toribio’s (2017) limitations with a follow-up study
- (ii) Scrutinize the divide between theoretical accounts and experimental findings

To accomplish these goals Frank and Toribio's (2017) aural repetition and sentence completion tasks are replicated with a Colombian Spanish group. The comparison of the new baseline data to the original Cuban test group facilitates a statistical analysis of between group effect, replacing speculation with definitive evidence. Given the Peninsular bias in the literature, if the assumption of limited variation in recomplementation proves to be false, then we can potentially reconcile the divide in theoretical accounts and experimental findings by proposing recomplementation as a locus of dialectal variation. We also consider the potential microvariations of not only intervener length, but also intervener type and sentence type. Analyzing microvariation from an experimental perspective permits a formal analysis of any potential interactions between group and type. The chapter continues with an overview of the linguistic phenomenon in question and a discussion of morphosyntactic variation in Cuban Spanish as compared to the Mainland baseline.

3.2 LINGUISTIC PHENOMENON

3.2.1 Syntactic-Theoretical Review

In the previous chapter we introduced the formal literature on recomplementation. Here we provide a more thorough review. Recomplementation has become a topic of frequent investigation in the formal syntactic literature grounded in Generativist and Minimalist programs (e.g., Brovotto, 2002; Demonte & Fernández-Soriano, 2009; Fontana, 1993; González i Planas, 2014; Gupton, 2010; Martín-González, 2002; Mascarenhas, 2007; Paoli, 2006; Radford, 2013; Villa-García, 2019, 2015, 2012). Utterances with multiple complementizers are attested in an array of statement and question examples as in (3.2) and (3.3), respectively. Additionally, a so-called quotative *que* can also precede a

secondary *que* within the same sentence (3.4) (for a review of quotative *que*, see Demonte & Fernández-Soriano, 2014; Extepare, 2010, Gonzalez i Planas, 2014). As shown, in these structures, a left-dislocated (LD) phrase is sandwiched between a high (C1) and a low (C2) complementizer.

- (3.2) Me dice que₁ por suerte, que₂ va a tener suficiente tiempo.
 CL₁SgDat say₃SgPres that for luck that is₃SgPres going to have enough time
 ‘S/he says that luckily s/he is going to have enough time.’
- (3.3) Me pregunta que₁ esa camisa, que₂ cuándo la iba a devolver.
 CL₁SgDat ask₃SgPres that that shirt that when CL₃SgFemAcc is₃SgPres going to return
 S/he asks when I am going to return that shirt.’
- (3.4) Que_{Quotative} si llueve, **que** no vienen. (Villa-García, 2015, 29)
 that if rains that not come
 ‘Somebody said/says that they won’t come if it rains.’

As the cartography of the left periphery has been developed and scrutinized, various accounts of the representation of recomplementation and function of the secondary complementizer have been put forth. These accounts are reviewed in detail in Villa-García (2015) and replicated in (3.5a-h) for the reader’s convenience:

- (3.5) a. *CP RECURSION* (e.g., Fontana, 1993, Iatridou & Kroch, 1992)
 [CP [C' que [CP Left-dislocate [C' que ...]]]]
- b. *FP* (Uriagereka, 1995)
 [CP [C' que [FP Left-dislocate [F' que ...]]]]
- c. *FINITENESSP* (e.g., Brovotto, 2002)
 [ForceP [For' que [TopicP Left-dislocate [Top' ... [FinitenessP [Fin' que ...]]]]]]
- d. *NO TOPICP/FOCUSP (Modified FINITENESSP)* (López, 2009)
 [ForceP [For' que [FinitenessP Left-dislocate [Fin' que ...]]]]
- e. *DISCOURSE PROJECTIONS* (Kempchinsky, 2013)
 [ForceP [For' que [DP Left-dislocate [D' que ... [FinitenessP [Fin' ...]]]]]]

In the FinitenessP account (3.5c), C1 heads ForceP and C2 heads FinitenessP. This account has been adopted by several theoreticians (e.g., Brovotto, 2002, López, 2009, Demonte & Fernández-Soriano, 2009). The No Topic/FocusP proposal (3.5d), a modified FinitenessP account, maintains the upper (ForceP) and lower (FinitenessP) bounds of the CP but does away with the topic-focus field, which Rizzi (1997) proposes must be activated by a topicalized and/or focalized constituent. The modified FinitenessP account assumes a FinitenessP with multiple specifiers in place of a CP that splits into a number of functional phrases. As in the original FinitenessP account, C1 heads ForceP and C2 heads FinitenessP. As Villa-García (2015) points out, one reason that these accounts are inadequate is they also fail “... to capture the fact that focalized material and interrogative complementizers must follow secondary *que*... unless we assume that such constituents are material in the TP domain” (63-64).

The DiscourseP proposal (3.5e) is viewed more favorably by Villa-García (2015). As in the modified FinitenessP proposal, this account discards TopicP and FocusP. However, this account incorporates a projection between ForceP and FinitenessP—the DiscourseP (DP), where C1 heads ForceP and C2 heads DP. This analysis accounts for the relationship between the left-dislocated element and the C2, where the latter is contingent on the former. Given this contingency, it is not surprising that both the sandwiched element and C2 are placed in [Spec, DP] and [Head, DP], respectively. Still, Villa-García (2015) concludes, “...a more complex structure (i.e., Rizzi’s detailed syntactic map) seems to be required in order to account for the complex range of phenomena that can occur in the Spanish left periphery” (66).

The so-called Moving Complementizers account (3.5f) adopts Rizzi’s proposed functional projections and proposes that C2 originates in FinitenessP and moves across the left-periphery. C1 and overt C2 is then the pronunciation of different copies (see Copy

Theory of Movement, Chomsky, 1995). Overt C2 is said to be predicted by the heaviness of the left-dislocated material. In our view, this is an attractive proposal given the experimental literature that supports length of intervening material as a probabilistic predictor of C2 lexicalization (Casasanto & Sag, 2008; Echeverría & López Seoane, 2019). Villa-García (2015) offers the counterexample (*Ya le dije que yo, que no voy* ‘I’ve already told him/her that I won’t go’), as well as several other refuting arguments. In our view, the occurrence of counterexamples is compatible with a probabilistic rather than categorical interpretation of secondary *que* expression.

Thus far we have reviewed pre-split CP theories (i.e., CP Recursion and FP), as well as theories that were proposed post Rizzi (1997) and considered functional projections to varying degrees (e.g., FinitenessP, modified FinitenessP, Discourse Projection, and Moving Complementizers). The last two accounts considered in the review are Doubled-Force (3.5g) (e.g., Demonte & Fernández-Soriano, 2009; Gupton, 2010; Martín-González, 2002) and TopicP (3.5h) (e.g., Paoli, 2006; Rodríguez-Ramalle, 2003; Villa-García, 2019, 2015, 2012). As explained in the previous chapter, these are two of the more widely adopted accounts. In the former account, C1 heads ForceP and C2 heads doubled-ForceP. Demonte and Fernández-Soriano (2009) specifically make the case that C2 is contingent on C1 and that the function of C2 can be to reintroduce the force of the sentence. Thus, it is not surprising that the complementizers head the same projection and reinforced projection, respectively.

Villa-García (2015) adopts the TopicP account. As in the Discourse projection account, C1 heads ForceP and C2 and the sandwiched element fill the head and specifier positions of the same phrase. The specific phrase is TopicP, whereas in the Discourse Projection proposal, it is DP. Villa-García considers TopicP account to be “empirically superior to the alternatives... as it straightforwardly accounts for the facts under discussion

without additional stipulations” (70). He cites not only the contingency of the C2 on the left-dislocated element, but also ellipsis licensing facts, the compatibility of the analysis with the possibility of more than one dislocated phrase and more than two complementizers, the ability for a left-dislocated phrase to occur below C2, the placement of the focalized or *wh*-elements in FocusP, the already understood recursive nature of TopicP without the need to propose recursion for any other projections, among other justifications with an abundance of cross-linguistic evidence.

3.2.2 The Function of Secondary *Que*

As chronicled in Villa-García (2019), the multifunctional nature of C2 goes beyond reintroducing force and marking topic. The “polyvalent” secondary *que* can also serve as a processing marker and a discourse marker, even possessing multiple roles at once. In terms of processing benefit, length of intervening material as a predictor of C2 lexicalization has been found in a couple of studies (Casasanto & Sag, 2008; Echeverría & López Seoane, 2019). It has further been alluded to in the syntactic-theoretical literature (Ledgeway, 2000; Radford, 2018; Villa-García, 2019). Casasanto and Sag (2008) interpret this result from a psycholinguistic, distance-based theory of linguistic complexity, where the integration cost associated with two elements engaged in a syntactic dependency is equal to the distance between them (Gibson, 1998, 2000; Gibson & Grodner, 2005). For example, when C2 is spelled out, the distance between the complementizer and the complement is zero. Thus, the cost of information retrieval is also zero. When C2 is null, C1 retrieval cost increases as the length of left-dislocated material increases. Another processing benefit is explained by Surprisal Theory. The more surprising (i.e., the less expected) the linguistic input, the more difficult it will be to process and the higher the

integration cost (for a review see Hale, 2001; Levy, 2008). Secondary *que* indicates that the left-dislocated segment has come to an end, reducing the processing cost associated with unexpected material or integration of the complement.⁴

Villa-García (2019) argues that C2, among its many functions, can also serve as a discourse marker. He cites several pieces of evidence that demonstrate how discourse markers and secondary *que* behave similarly. For example, “discourse markers signal a relationship between the element they introduce (i.e., the embedded clause) and the prior segment (i.e., the dislocated material)” (31). Furthermore, similar to the processing marker, discourse particles and the C2 can facilitate processing for the hearer (also see Hansen, 1998). Also, in line with the syntactic-theoretical accounts in the previous section, discourse markers and the C2 are grammatically optional. Villa-García offers a long list of similarities between the two phenomena, which include but are not limited to a clustering effect (e.g., *pues que*), a tendency to be spoken rather than written (although see Echeverría & López Seoane, 2019), and their derivation from conjunctions (e.g., Fraser, 1999).

3.2.3 The Disparity between Theoretical Accounts and Experimental Findings

As reviewed in Section 3.2.1, syntactic-theoretical literature largely aligns on the grammaticality and optionality of overt C2 in present-day Spanish. Further, Fontana (1993) and Echeverría and López Seoane (2019) document widespread use of recomplementation in their analysis of old written Spanish, arguing that usage in old Spanish was more frequent

⁴ As noted in Frank and Toribio (2017), “single complementizers are more likely to be lexicalized in English when the complement clause is less predictable, such as when there is material intervening between the complement-taking verb and the subordinate clause. Thus, there is a processing benefit that is accrued by lexicalizing the complementizer in English, whether in single or multiple complementizer sentences (Jaeger, 2006, 2010)” (122).

than in present day and the null C2 variety was the marked form.⁵ Importantly, claims of widespread use in present-day unguarded spoken speech are not restricted to Spanish. Rather, they include Portuguese (e.g., Mascarenhas, 2007), Catalan (González i Planas, 2014), Galician (e.g., Gupton 2010), Italian (e.g., Paoli, 2006; Rizzi, 2013) and English (e.g., Radford, 2018, 2013).

In the previous chapter, we reviewed the existing experimental literature on recomplementation and found that overt C2 was either marginally acceptable and associated with a decrement in acceptability judgment (Frank, 2016), ungrammatical (Casasanto & Sag, 2008) and/or probabilistically predicted by an extralinguistic variable, e.g., intervener length or strain on processing and memory (Casasanto & Sag, 2008; Echeverría & López Seoane, 2019). As referenced in Section 3.2.2, syntactic-theoretical accounts have also hinted at an effect of length or “heaviness” in C2 lexicalization patterns. For example, Radford’s (2018) English corpus displays on average 5.9 words per phrase that is flanked by a complementizer; this length closely matches what researchers who adopt formal experimental methods define as “long” intervening material. Further, Ledgeway’s (2000) Moving Complementizers formal account notes that secondary *que* is more likely to be pronounced when the dislocated material is heavy.

In summary, we cannot rule out that a processing-based proposal may adequately account for the probabilistic (not optional) occurrence of secondary *que* (e.g., a processing/grammaticality tradeoff theory; see Casasanto & Sag, 2008). As noted in Frank and Toribio (2017), “...processing explanations are proving increasingly promising as researchers reevaluate syntactic phenomena that were formerly understood in grammatical terms...” (122) (e.g., O’Grady, 2010). We add that processing explanations are proving

⁵ Importantly, Fontana (1993) does offer a word of caution that the overt variety is restricted to fewer contexts in modern Spanish.

promising for the reevaluation of phenomena previously understood as disfluencies, errors and/or *ungrammatical* (e.g., Casasanto & Sag, 2008; Ferreira & Patson; Ferreira et al., 2004; Polinsky & Scontras, 2020). We hold off on a deeper discussion of the topic until the Chapter 4.

Importantly, both sides of the syntactic-theoretical and experimental debate assume that secondary *que* is produced in everyday discourse. A usage-based account of the phenomenon may then not find the question of grammaticality all that interesting (e.g. Echeverría & López Seoane, 2019). In our view, the apparent contradiction between the findings deserves further scrutiny because it sheds light on a broader discussion. Namely, it raises questions around the replicability of empirical research and the effect of different methodologies. It also highlights the aforementioned standard variety research bias and the importance of analyzing potential microvariations within a phenomena.

3.3 MORPHOSYNTACTIC VARIATION IN SPANISH

3.3.1 Cuban and “Mainland” Spanish

The present chapter sheds light on the gap between theoretical accounts and experimental evidence by investigating the phenomenon of recomplementation and its potential microvariations in Cuban and Colombian Spanish. Given the Peninsular bias in the literature, we specifically investigate these two language varieties to shed light on whether the assumption of limited variation in recomplementation is true. Cuban Spanish shares many linguistic features with its Caribbean counterparts and other regions along the Atlantic coast. These features also distinguish said varieties from “Mainland” ones, such as non-coastal Colombian Spanish, which we adopt as a proxy for Mainland Spanish in the present chapter (e.g., Otheguy & Zentella, 2012; Otheguy et al., 2007). This is exemplified

most recently by the timely compendium *Cuban Spanish Dialectology: Variation, Contact and Change* (Cuza, 2017), which offered novel data on morphosyntactic, lexical, and phonological and phonetic features across multiple generations and regions. In this section, we review the work on morphosyntactic variation in the volume, motivating the selection of our test group and baseline.

With respect to morphosyntactic features specifically, two salient dialectal features are represented in the collected contributions: word order and pronoun expression. Erker et al. (2017) researched subject-verb/verb-subject (SV/VS) word order in adult Cubans residing in NYC. Participants were divided into two groups— those who had lived in NY for less than five years, “newcomers” (1st generation), and those who were either born and raised in NY or had moved before the age of 8, “NY-raised” (2nd generation).⁶ The former group were assumed to behave similarly to Cubans currently residing in Cuba. The aim of the study was to shed light on the factors that predict word ordering and to explore the suggestion that Cubans, along with the Caribbean lect more generally, prepose subjects at a higher rate than the rest of the Spanish-speaking world. Of the compiled 700-item corpus, 84% and 88% of the items presented SV word order in newcomers and NY-residents, respectively. This word order was significantly predicted by verb type, subject type, sentence type, subject referent and clause type variables. Even $\frac{2}{3}$ of the interrogative tokens demonstrated a preposed subject (e.g., *Dónde Juan quiere comer* ‘Where does John want to eat’), where the post position (i.e., subject-verb inversion) is more common in the Mainland variety. The authors conclude that while some differences were found between newcomer and NY-raised groups, their behavior is qualitatively the same.⁷ The strong

⁶ Language use and proficiency contrasted between groups, with newcomers demonstrating higher rate of Spanish use and a higher proficiency overall.

⁷ NY-raised postpose less than newcomers, clause type only reaches significance for NY-raised, and there appears to be a general tendency to reduce word order flexibility in situations of contact.

preference for preverbal subject position across groups is interpreted as cross-generational stability, as children acquire the language of their parents.

Ortiz-López et al. (2017) also investigate SV/VS word order, along with other interrelated Null Subject Parameter (NSP) properties (e.g., pronominal expression). As the authors note, Spanish, likely due to a Mainland bias, is traditionally considered “... a pro-drop language, with free variation in terms of subject-verb order (SV/VS)....” (98). The authors investigate these two features in Cubans residing in Havana via naturalistic data (i.e., 30-minute semi-spontaneous interviews). SV/VS results suggest that while the vast majority of items have a preposed subject, certain conditions are more categorical than others. For example, within the factor of subject type, 97% of pronominal subjects were preverbal (e.g., 99% of second person pronoun items), while a vast majority but more variable percentage of nominal subjects were preverbal (81%). With respect to syntactic complexity and clause type factors, complex NPs (66%) and relative clauses (78%) were the most variable conditions, respectively. The variability of word order in relative clauses supports previous research on Caribbean Spanish and points to the importance of investigating the microvariation within a given linguistic phenomenon (e.g., Otheguy & Zentella, 2012).

In terms of rate of pronominal expression, Ortiz-López et al. (2017) found that 30% of the tokens in the coded semi-spontaneous interview expressed the pronoun, with second person singular (*tú* ‘you’) and first person singular (*yo* ‘I’), appearing nearly fixed in the preverbal position (also see Ortiz-López, 2009; Otheguy & Zentella, 2012).⁸ The 30% expression rate is higher than the Mainland varieties (Ecuadorians, Colombians, Mexicans of New York City, Spaniards, Mexicans of Mexico City), which present an average

⁸ Though infrequent in the corpus, *uno* ‘one’ was also frequently spelled out.

production rate of 23% (Lastra & Butragueño, 2015; Otheguy & Zentella, 2012). The authors argue that their finding supports a more general contrast between the Caribbean lect and Mainland Spanish (e.g., Aponte Alequín & Ortiz-López, 2015; Bosque et al., 1999; Bosque & Gutiérrez-Rexach, 2009).

In a second study on pronominal expression, Alfaraz (2017) investigated 57 Cubans born between 1885 and 1970 via existing corpora. 8300 tokens were analyzed for significant effects between pronominal expression and several linguistic and social variables across three distinct age groups. Whereas Ortiz-López et al. found a 70/30 null vs. overt pronoun split, this study returned a 65/35 split. Moreover, as noted by the author, their finding of a 35% rate of pronominal expression is remarkably similar to Otheguy et al.'s (2007) 33% pronominal expression in newcomers and Otheguy and Zentella's (2012) 38% pronominal expression for varying lengths of residence. This motivates their conclusion that pronominal expression is stable in the Cuban Spanish variety. Alfaraz specifically finds that person-number (i.e., 2nd person singular and 1st person singular) then coreference (i.e., switch context) factors most strongly predicted pronominal expression. The author explains that the switch reference effect has been demonstrated to be robust across Caribbean (e.g., Alfaraz 2015; Otheguy & Zentella, 2012; Otheguy et al., 2007) and non-Caribbean varieties (e.g., de Prada Pérez, 2009; Silva-Corvalán, 1994; Otheguy et al., 2007). Alfaraz's comparative analysis across birth year demonstrates that person-number and coreference factors remain significant across time, adding credence to the stability of pronominal expression in Cuban Spanish.⁹

In sum, this review highlights areas of morphosyntactic variation in Cuban Spanish (and the Caribbean lect) when compared to Mainland Spanish. Importantly,

⁹ Weaker effects of verb type, tense mood aspect, and polarity do seem to interact with birth year. For example, the TMA predictor weakened in each progressive age group, a finding that should be further scrutinized in future studies.

microvariations within the general phenomena were also considered. For example, Cuban Spanish displays a more rigid SV word order, which is nearly fixed for pronominals though less so for nominals, and more rigid in relative clauses than complex NPs. Subject lexicalization also occurs at a higher rate in Cuban Spanish and is most frequent in 2nd and 1st person singular. Several commonalities between the two groups were also highlighted in this review. For example, despite claims of free variation of SV-VS word order in Mainland Spanish, SV occurs at a much higher rate, much like in Cuban Spanish. Also, pronominal expression is predicted by switch referent contexts in both groups. Frank & Toribio (2017), the study that we follow up in this chapter, sought to contribute to Cuza's (2017) compendium on Cuban Spanish by investigating recomplementation as an unexplored phenomenon and potential locus of morphosyntactic variation. While a deeper discussion of the methods and results will be discussed later, we offer a brief summary of the study in Section 3.3.2.

3.3.2 Frank and Toribio (2017)

Frank & Toribio (2017) were motivated to investigate recomplementation in Cuban Spanish for several reasons. One reason was to investigate a potential dialectal bias in the syntactic-theoretical literature. As noted earlier, the latest volume on the structure of multiple *que* sentences draws primarily from Peninsular/European data (Villa-García, 2015). This selection was further motivated by the existing literature demonstrating varying patterns of *que* expression more generally across the Spanishes of the world. As reviewed in the previous chapter, Silva-Corvalán (1993) investigated patterns of *que* expression in Los Angeles heritage speakers of Spanish. She concluded that observed null *que* in argument clauses is the result of cross-linguistic influence effects in the direction of

the minority language conditioned by surface overlap (e.g., Müller and Hulk, 2001; Yip and Matthews, 2009). More recently, Riccelli (2018) searched Twitter for instances of *que*-drop in Mexico City and LA. Curiously, he found *que*-drop to be a more prevalent phenomenon in the former population. This finding brings into question Silva-Corvalán's (1993) claim of *que*-drop being motivated by contact with English. Further, as discussed in Frank and Toribio (2017), "...Riccelli (p.c.) explains that while instances of *que*-drop are abundant in Mexico City tweets, they are almost non-existent in Peninsular Spanish tweets" (123). Lastly, Cuza and Frank (2011/2015) investigated heritage speakers' and second language learners' representation and usage of primary *que* (C1) in indirect questions. Coincidentally, they found several instances of *que*-drop in the Cuban-speaker baseline in contexts where overt *que* is thought to be obligatory (e.g., indirect questions introduced by non-ask/wonder verbs). Together, these studies add credence to the argument that *que*-drop patterns are variable in Spanish.

With these motivations in mind, Frank and Toribio (2017) assessed the grammaticality and potential processing benefits of secondary *que* in Cuban Spanish via two controlled offline production tasks: aural repetition and sentence completion. As explained in Frank and Toribio (2017), when repeating an aural prompt in a repetition task, it is assumed that "...constructions that are consistent with the speakers' grammars will be faithfully reproduced, while those that are not will be distorted or manipulated during production (e.g., Hamayan et al., 1978; Pérez-Leroux et al., 2011; Radloff & Hallberg, 1991)" (125). The sentence completion task was adapted from Cuza and Frank's (2011) production task, which successfully elicited high *que*.

Unlike in the previous chapter, Frank and Toribio (2017) operationalize intervener length as a fixed effect, given the literature that finds length is a significant predictor of C2 lexicalization (e.g., Casasanto & Sag, 2008; Echeverría and López Seoane, 2019). Data

elicited from 25 participants submitted to statistical analyses indicated that reduplicative *que* is dispreferred in this dialect of Spanish, independent of intervener length. Specifically, null reduplicative *que* conditions were repeated with greater accuracy than those with lexicalized reduplicative *que*. This result was interpreted as suggesting that multiple complementizers are not licensed by the grammar of Cuban Spanish. Moreover, only 1% of the test items displayed reduplicative *que* insertion in the sentence completion task, indicating that complementizer doubling does not serve a facilitative function in the offline production tasks. These findings contrast with the extensive syntactic-theoretical literature on recomplementation in Peninsular Spanish, as reviewed in Section 3.2.1

Importantly, one limitation of Frank and Toribio (2017) is the lack of a Mainland Spanish experimental group. To our knowledge, this is the first study to investigate recomplementation via aural repetition and sentence completion methods. Thus, to ensure maximal comparison, these methods should be replicated with the comparison group. Further, an analysis of the linguistic restrictions that govern C2 expression is limited to the length of intervener factor. It will also be important to investigate the potential for contrastive behaviors of the type of left-dislocated element and sentence. The addition of a between group comparison and more in-depth analysis of potential linguistic microvariations are accounted for in the present chapter.

3.3.3 Research Questions

The present chapter addresses the limitations to Frank and Toribio (2017) with a follow-up study that replicates the original study of recomplementation in Cuban (Caribbean) Spanish with a new, Colombian (Mainland) test group. The study reprises the research questions from Frank and Toribio:

- (RQ 1) Is overt C2 licensed by Colombian ‘Mainland’ Spanish grammar unlike in Cuban?
- (RQ 2) Is recomplementation a locus of dialectal variation, microvariations considered?
- (RQ 3) Does overt C2 facilitate complement integration (e.g., does intervener length predict secondary *que* expression)?

In so doing, it further scrutinizes the grammaticality of lexicalized secondary *que* in recomplementation structures, with specific reference to the gap between theoretical accounts and experimental findings. With respect to (RQ1), we anticipate that the Mainland Spanish group will accurately repeat sentences with a secondary *que* in the overt C2 condition. Given the assumption of the aural repetition task that constructions that are consistent with the speakers’ grammars will be faithfully reproduced, we predict that secondary *que* expression is licensed by Mainland Spanish grammar. This prediction is supported by both experimental evidence and syntactic-theoretical accounts. In the previous chapter, the Colombian baseline displayed a decrement in acceptability judgment in the overt C2 condition that was still within the range of marginal acceptability (i.e., 2.82 and 2.99 on a 1-totally acceptable to 7-totally unacceptable scale). This result is made all the more intriguing when we consider intervener length was held constant in the short condition, which is not thought to predict C2 lexicalization (Casasanto & Sag, 2008; Echeverría & López Seoane, 2019). Secondary *que* is also supported the syntactic-theoretical literature that argues for the grammaticality of secondary *que*, most likely, from the perspective of a Mainland bias. Importantly, any secondary *que* expression in the null C2 condition of the repetition task or in the sentence completion task, while not direct evidence, would support our expectation.

For (RQ2), we weigh our prediction of Mainland Spanish in (RQ1) against Frank and Toribio’s (2017) preliminary conclusions that overt C2 is not licensed by Cuban

Spanish grammar. Thus, we anticipate that recomplementation is indeed an unexplored locus of dialectal variation. With respect to the aural repetition task, we specifically predict that the overt secondary *que* condition will be accurately reproduced at a significantly higher rate in the Mainland Spanish baseline when compared to the Cuban test group. Any interactions between group and the microvariation variables of intervener length, intervener type and sentence type will also support the prediction. Further support can be found in the divergence between groups with respect to the insertion of *que* in the null C2 condition of the repetition task or in the sentence completion task. Recomplementation as a locus of variation would help to explain the seeming contradiction between experimental and syntactic-theoretical evidence.

Lastly, with respect to (RQ3), previous experimental studies have found that overt C2 is more likely to be lexicalized when the intervener length is long, i.e., four words or more (Casasanto & Sag, 2008; Echeverría & López Seoane, 2019). This evidence is supported by the syntactic theoretical literature (Ledgeway, 2000; Radford, 2018; Villagarcía, 2019). Curiously, Frank & Toribio (2017) did not find an effect for length. We predict that intervener length is a locus of dialectal variation and anticipate it will significantly predict C2 lexicalization in the Mainland baseline. Given a scarcity of evidence, we do not expect sentence type (i.e., embedded questions versus statements) or intervener type (direct/indirect objects, adverbs/adverbials) to reach significance. If the sentence type does, then the results from the previous chapter on heritage Spanish will be further validated. We cannot think of a plausible explanation for why intervener type might reach significance.

3.4 THE STUDY

3.4.1 Participants

In order to test these predictions, 16 participants were recruited from Medellín, Colombia, to serve as a proxy for Mainland Spanish. They were compared to Frank & Toribio's (2017) Cuban Spanish group (n=25), originally recruited in Havana, Cuba. The Colombian group were recruited through word of mouth with the support of a local Spanish instructor who was trained as a research assistant. All demographic information was solicited via language background questionnaire (see Appendix A). Selected metadata is summarized in Table 3.1, with Cuban data replicated from Frank and Toribio (2017) for the reader's convenience.

Table 3.1. Selected demographic information

Selected Information	Colombian Spanish (n=16)	Cuban Spanish (n=25)
Birthplace	Colombia	Cuba
Current Residence	Medellín	Havana
First language	Spanish	Spanish
Parent's First Language	Spanish	Spanish
Gender	6 male, 10 female	11 male, 14 female
Age at testing	18-53 ($M= 31$)	18-78 ($M= 34$)
Level of Education	10 university 4 technical school 2 high school	18 university 4 technical school 3 high school

All Mainland Spanish participants were born in Colombia and currently residing in the city of Medellín. In all cases their first language, as well as the first language of both their parents was Spanish. Their ages ranged from 18-53 ($M= 31$) and they held such occupations as student, web developer, unemployed, house parent, and engineer, among others. Lastly, very little variation was found in terms of level of education. All had

received at minimum a high school level of education, with the vast majority having attended post-secondary schooling (14/16). Similarly, Frank & Toribio’s Cuban participant group reported Spanish as both their first language and the first language of their parents. These participants were currently residing in Havana at time of testing and ranged in age from 18-78 ($M= 34$). All participants had received at minimum a high school level of education (22/25 post-secondary) and held such occupations as student, teacher, engineer, accountant, journalist, physical therapist, army employee and retiree, among others. A summary of each group’s language history and contact with English is included in Table 3.2.

Table 3.2. Language history, use and contact

Selected Information	Colombian Spanish (n=16)	Cuban Spanish (n=25)
<i>Language of Instruction</i>		
• elementary school	Spanish	Spanish
• high school	Spanish	Spanish
• higher education	Spanish	Spanish
<i>Spanish Language Use</i>		
• at school	mainly Spanish	mainly Spanish
• at home	Spanish only	Spanish only
• at work	mainly Spanish	mainly Spanish
• in social situations	Spanish only	mainly Spanish
Frequency of visits to English-speaking countries	never (10)	never (23)
	almost never (2)	almost never (1)
	infrequently (4)	infrequently (1)

As demonstrated in Table 3.2, both the Mainland Spanish and the Cuban Spanish groups’ primary language of instruction was Spanish. This is true from elementary school through levels of higher education. Furthermore, current language use across the contexts of school (if relevant), home, work and social situations was calculated on a scale of 0

“English only”, 1 “mainly English”, 2 “a little more English”, 3 “both equally”, 4 “a little more Spanish”, 5 “mainly Spanish”, to 6 “Spanish only”. On average, both groups used mainly Spanish or Spanish only across all contexts. Specifically, 16 out of 16 baseline participants reported that they used only or mainly Spanish in social contexts, and 14/16 used only or mainly Spanish at work, while the remaining two used slightly more Spanish and Spanish and English equally, respectively. Similarly, 22 out of 25 participants reported that they used only or mainly Spanish both at work and in social contexts, while the remaining three used slightly more Spanish. Lastly, the vast majority of the participants had never traveled to an English-speaking country, 10/16 and 23/25 in the mainland and Cuban groups respectively. Of those who had, visits were either infrequent (e.g., “every two years”) or almost never (e.g., “once for 15 days”). In summary, very little variation was found among the participants in terms of language use, both historically and at present.

3.4.2 Methods and Design

In order to investigate recomplementation as a potential locus of dialectal variation and to further scrutinize the incongruency between experimental and syntactic-theoretical evidence we replicated Frank and Toribio’s (2017) two offline aural production tasks and analyzed independent variation, as summarized in Table 3.3.

Table 3.3. Variables under investigation

Independent Variables	Levels	Elicited Repetition	Sentence Completion
Group	Cuban Spanish	✓	✓
	Mainland Spanish		
Secondary <i>que</i>	Null	✓	×
	Overt		
Intervener length	Short (2 words)	✓	✓
	Long (7 words)		
Intervener type	Direct object	✓	✓
	Indirect object		
	Adverbials		
Sentence type	Question	✓	✓
	Statement		
Dependent Measure		Repetition accuracy (binary)	Secondary <i>que</i> insertion (binary)

The group variable divides into two levels, including Frank and Toribio’s Cuban Spanish sample and our Colombian Spanish group, which served as a proxy for Mainland Spanish. Secondary *que*, which only served as a variable for the elicited repetition task, manipulated the conditions of null versus overt C2. Intervener length refers to the number of words of the dislocated material. The short condition was fixed at two words, and the long condition was fixed at seven. Intervener type, on the other hand refers to the category of material. For this purpose, we adopted three levels, direct objects (e.g., *esa camisa*), indirect objects (e.g., *al doctor*) and adverbs or adverbial clauses (e.g., *por supuesto*). The direct and indirect object levels were collapsed in the second task. Lastly, sentence type pertains to whether the item was an indirect question or an assertion, where all questions

were introduced by the matrix verb *preguntar* ‘to ask’ and all statements were introduced by the matrix verb *decir* ‘to say/tell’, see test item examples below.

With respect to the two experimental tasks, we refer to Frank & Toribio’s (2017) description. The elicited repetition or imitation task and the sentence completion task are both designed to elicit oral secondary *que* expression. We specifically adopted speaking and listening modes as opposed to reading and writing modes to offer a more natural context for recomplementation on present-day Spanish. In the elicited repetition task, participants were instructed to listen to a sentence one time and then repeat it aloud. As explained in Frank and Toribio (2017), the assumption in an elicited repetition task is “...constructions that are consistent with the speakers’ grammars will be faithfully reproduced whereas those that are not will be altered” (Hamayan et al., 1978; Pérez-Leroux, Cuza & Thomas, 2011) (125). Further, we examine “...the alternative possibility that a secondary complementizer will be employed as a strategy for integrating complement clauses when required by a long intervener” (125). See examples (3.8ab) borrowed from Frank and Toribio (2017):

(3.8) *Elicited imitation*

- a. Me dice que sin duda (que) va a haber mucho daño al techo después de la tormenta severa.
‘S/he tells me that without doubt (that) there will be significant damage to the roof after the severe storm.’
- b. Me dice que sin duda después de la tormenta severa (que) va a haber mucho daño al techo.
‘S/he tells me that without doubt after the severe storm (that) there will be significant damage to the roof.’

As demonstrated in (3.8a) and (3.8b), the statements introduced by the verb *decir* are of equivalent overall length. Sentence length is an important element of the design of repetition tasks, where the strain on memory should be significant but not so much so as to

interfere with task execution. After piloting for length, we aligned on a fixed 17-word count (or 18 words in the overt C2 condition). Importantly, (3.8a) and (3.8b) differ in intervener length, two words and seven words, respectively. These adverb/adverbial interveners are flanked by a primary and secondary *que*. The secondary *que* is placed in parentheses here to represent to the reader that half of the test items include an overt C2 and half include a null C2. A total of 24 test items composed of varying conditions of secondary *que*, intervener length, intervener type and sentence type were created in total and scrambled with 36 distractors of comparable length, targeting unrelated subject-verb inversion and pronominal expression phenomena. See Appendix E for the full list of experimental stimuli.

According to Frank & Toribio (2017), for task two, participants first listened to a preamble that contextualized the test item. Then they heard a prompt for an incomplete sentence. Their task was to repeat the prompt and complete the sentence. As demonstrated in example (3.9), prompts included either short or long topicalized material.

(3.9) *Oral sentence completion*

Preamble: Si ganas (la competencia de arte anual), ¿cómo vas a gastar el premio monetario?
'If you win (the annual art competition), how will you spend the monetary prize?'

Prompt: Me pregunta que si gano (la competencia de arte anual)...
'S/he asks me that if I win (the annual art competition)...'

Expected Response: Me pregunta que (si gano la competencia de arte anual) ...
(que) cómo voy a gastar el premio monetario.
'He asks me that if I will the annual art competition (that) how I will spend the monetary prize.'

Example (3.9), represents a question condition test item with an adverbial clause intervener type and both short (and long) conditions. The participant's task is to first listen to the preamble followed by the prompt. As explained by Frank & Toribio (2017), all

prompts were created by first recording a native speaker saying the entire expected response. The recording was then edited in Audacity® software (Audacity Team, 2019) by deleting the complement. This resulted in a prompt with an authentic continuation rise prosody (Cuza & Frank, 2015). After listening to the prompt, the participant first repeated the prompt then completed the sentence in order to produce their own complete sentence. A total of 16 test items composed of varying intervener length, intervener type and sentence type conditions were created in total and scrambled with 16 distractors of comparable length. The only independent variable from the repetition task that wasn't investigated here is secondary *que*. This is because secondary *que* expression is actually what is being measured in the sentence completion task. See Appendix F for the full list of experimental stimuli.

3.4.3 Results

As in Toribio and Frank (2017), we code repetition task data with a score of 1 for accurate imitation and a score of 0 for inaccurate imitation. Inaccurate imitation was defined as “*que* insertion or omission in the null and overt condition, respectively...” (126). Further, secondary *que* aside, “...only sentences that were imitated with a maximum of two word omissions, alterations, or insertions were included in the final analysis” (126). This served as a proxy for attention to task and resulted in the discarding of 49 items out of a 600 item corpus (8%) for Cuban Spanish and 39 items out of a 384 items corpus (10%) for Colombian Spanish .

Importantly, for the purpose of this chapter we offer a novel analysis of the data. Not only have we included the Colombian group in order to replace speculation of dialectal variation with definitive evidence but we consider the potential microvariations of not only

intervener length, but also intervener type and sentence type. Descriptive results for experiment one are summarized for the Cuban and Colombian groups in Table 3.4 and Table 3.5, respectively.

Table 3.4. Elicited imitation descriptive results for Cuban Spanish

Variables	Levels	Target Imitation	Nontarget imitation	Proportion target
Secondary <i>que</i>	Null	167	10	0.94
	Overt	70	204	0.26
Intervener length	Short (2 words)	163	107	0.60
	Long (7 words)	174	107	0.62
Intervener type	Direct object	121	65	0.65
	Indirect object	105	71	0.60
	Adverbials	111	76	0.59
Sentence type	Question	176	97	0.64
	Statement	161	115	0.58

Table 3.5. Elicited imitation descriptive results for Colombian Spanish

Variables	Levels	Target Imitation	Nontarget imitation	Proportion target
Secondary <i>que</i>	Null	153	1	0.99
	Overt	37	121	0.23
Intervener length	Short (2 words)	74	65	0.53
	Long (7 words)	116	57	0.67
Intervener type	Direct object	79	44	0.64
	Indirect object	51	29	0.64
	Adverbials	73	48	0.60
Sentence type	Question	108	57	0.65
	Statement	95	65	0.59

As demonstrated in Table 3.4, the Cuban group has a much higher proportion of target responses in the null C2 (0.94) than the overt C2 (0.26) condition of the secondary *que* variable. In other words, when the participant is asked to repeat sentence items with the null C2, they do so very accurately, without inserting the *que*. On the other hand, when the participant is asked to repeat sentence items with the overt C2, they do so very inaccurately, by removing the *que*. There does not appear to be clear effect for the other variables as there is for secondary *que*. That is to say, for intervener length, short lengths and long lengths are repeated with very similar accuracy. Similarly, the levels within intervener and sentence type are also repeated with similar accuracy. In Table 3.5, we see that the Colombian group demonstrates very similar behavior, with a much higher proportion of target responses in the null C2 (0.99) than the overt C2 (0.23) condition of the secondary *que* variable. Further there is no clear effect for the other variables. The similar between-groups behavior is highlighted in Table 3.6, where the proportion target responses for each group across all the variables and levels therein, as well as the overall proportion mean, is laid out side by side.

Table 3.6. Elicited imitation between groups comparison

Variables	Levels	Proportion target Cuban	Proportion target Colombian
Secondary que	Null	0.94	0.99
	Overt	0.26	0.23
Intervener length	Short (2 words)	0.60	0.53
	Long (7 words)	0.62	0.67
Intervener type	Direct object	0.65	0.64
	Indirect object	0.60	0.64
	Adverbials	0.59	0.60
Sentence type	Question	0.64	0.65
	Statement	0.58	0.59
	<i>Mean</i>	0.59	0.62

The Cuban and Colombian groups' accuracy in the repetition task is strikingly similar with their proportion of target responses within 0.10 (10%) of each other across every level. Further the overall mean of proportion target responses for Cuban and Colombian is 0.59 and 0.62, respectively. Importantly, the only variable that seems to be having an effect in both groups is secondary *que*.

In order to investigate these descriptive statistics and to shed light on research questions (1-3), a general linear mixed effects model was run with the GLMER function in R (R Core Team, 2017). The model defined one random intercept for subject, five fixed effects, group (Cuban, Colombian), secondary *que* (null, overt), intervener length (short, long), intervener type (DO, IO, adverbial), and sentence type (question or statement), as well as interactions between group and all other variables and secondary *que* and all other variables. The interactions with group were specifically defined to shed light on

recomplementation as a potential locus of dialectal variation, with microvariations considered. The interactions with secondary *que* investigate the facilitative nature of overt C2. For example, secondary *que**intervener length speaks to the documented length effect. The results of this analysis are summarized in Table 3.7.

Table 3.7. Elicited imitation statistical analysis

Statistical Design	Results
<i>Fixed Effects</i>	
• Group	$\beta = -1.938, z = -1.109, p = 0.267$
• Secondary <i>que</i>	$\beta = -6.589, z = -7.929, p < 0.001$
• Intervener length	$\beta = -0.039, z = -0.030, p = 0.976$
• Intervener type	$\beta = -0.753, z = -1.400, p = 0.161$
• Sentence type	$\beta = -0.598, z = -1.470, p = 0.141$
<i>Interactions</i>	
• Group*Second. <i>que</i>	$\beta = 1.252, z = 1.464, p = 0.143$
• Group*Inter. length	$\beta = -0.896, z = -1.776, p = 0.075$
• Group*Inter. type	$\beta = 0.315, z = 0.496, p = 0.620$
• Group*Sent. type	$\beta = 0.113, z = 0.230, p = 0.818$
• Second. <i>que</i> *Inter. length	$\beta = -0.598, z = -1.470, p = 0.141$
• Second. <i>que</i> *Inter. type	$\beta = 0.898, z = 0.476, p = 0.332$
• Second. <i>que</i> *Sent. type	$\beta = -0.774, z = -1.002, p = 0.197$

Secondary *que* was confirmed to be significant ($\beta = -6.589, z = -7.929, p < 0.001$), while no other variable reached significance. Furthermore, none of the group interactions or the secondary *que* interactions reached significance. In the absence of significant interactions, no post hoc analysis was run. However, in order to investigate the facilitatory

effect of overt C2 a bit further (RQ3), we analyzed the relationship of age (18-78, $M=34$) and repetition accuracy of overt C2 test items. The null C2 test items were not explored due to insufficient variation (i.e., ceiling performance). We hypothesized that if overt C2 served a facilitatory function, then accuracy would increase with age, independent of the grammatical status of C2 (see literature on normal cognitive aging, e.g., Harada et al., 2013). As demonstrated in Table 3.8, a Gaussian Mixture Model (GMM) found that the data points clustered in three distributions (Reynolds, 2009):

Table 3.8. Trimodal distribution of age and overt C2 repetition accuracy

Age Group	Mean Accuracy	Interpretation
18-32 (n=28)	0.299 (<i>SD</i> : 0.259)	Low accuracy (30%)
46-53 (n=10)	0.166 (<i>SD</i> : 0.178)	Very low accuracy (17%)
66-78 (n=3)	0.000 (<i>SD</i> : 0.000)	No accurate responses (0%)

The results of the data clustering analysis show an opposite pattern to what was hypothesized. Accuracy of overt C2 repetition actually decreases with age. This does not support a C2 facilitatory hypothesis. Secondary *que* appears to be treated as an extra element to remember (either optional or ungrammatical), which becomes increasingly more challenging with age. Alternatively, in the following chapter, we propose a model where facilitation effects associated with overt C2 benefit those with higher working memory capacity. Further, we cannot rule out the possibility that the age effect is driven by social factors rather than cognitive decline. Future studies can shed light on this empirical question.

In sum, the significant effect of secondary *que* tells us that participants repeat null C2 test items with significantly greater accuracy than overt C2 test items. This confirms

for Colombian Spanish what Frank and Toribio (2017) found for Cuban Spanish. Further, given the underlying assumption of the repetition task that constructions that are consistent with the speakers' grammars will be reproduced whereas those that are not will be remedied, we interpret this finding as evidence that overt C2 is not licensed by Mainland Colombian Spanish grammar (RQ1). The lack of effect for group and the absence of interactions between group and secondary *que*, intervener length, intervener type and sentence type provide robust evidence that recomplementation is not a locus of dialectal variation (RQ2). This evidence disconfirms the suggestion that secondary *que* might be an example of morphosyntactic variation in Cuban Spanish (Frank & Toribio, 2017). Further, between the absence of a significant interaction between intervener length and secondary *que*, the ceiling repetition accuracy of the null C2 condition, and the indirect relationship between age and repetition accuracy of the overt C2 condition, no evidence was found in favor of an overt C2 facilitatory effect (RQ3).

In order to supplement the repetition task and to further investigate whether recomplementation usage varies across dialects and/or whether C2 facilitates complement integration ([RQ2] and [RQ3], respectively), we replicate Frank and Toribio's (2017) sentence completion task. Recall that participants are instructed to first listen to a contextualizing preamble followed by an incomplete sentence prompt. They are then to repeat the prompt and complete the sentence. When participants inserted the secondary *que*, their response was coded with a value of 1. For instances of null C2, the response was coded with a 0. As in Toribio & Frank (2017), all items where the preamble was not repeated were discarded from the dataset. Importantly, the preamble repetition allows the researcher to distinguish between direct and indirect speech, where in the former, the conditions for C2 lexicalization would not be met. This along with incongruencies between

prompt and response resulted in 15 out of 400 responses (4%) being removed from the CS dataset and 2 out of 240 items (1%) being removed from the MS dataset.¹⁰

As in the elicited imitation task, we expand on Frank and Toribio's (2017) data analysis by investigating not only the Colombian group but also potential microvariations of intervener length, intervener type and sentence type. Descriptive results for experiment two are summarized for the Cuban and Colombian groups in Table 3.9 and Table 3.10, respectively.

Table 3.9. Sentence completion descriptive results for Cuban Spanish

Variables	Levels	<i>Que</i> insertion	<i>No que</i> insertion	Proportion <i>que</i>
Intervener length	Short (2 words)	1	192	0.005
	Long (7 words)	3	189	0.015
Intervener type	Objects	2	191	0.010
	Adverbials	2	190	0.010
Sentence type	Question	4	184	0.021
	Statement	0	197	0.000

¹⁰ In addition, one Colombian participant was discarded from the dataset, since they categorically did not repeat the prompt before answering.

Table 3.10. Sentence completion descriptive results for Colombian Spanish

Variables	Levels	<i>Que</i> insertion	No <i>que</i> insertion	Proportion <i>que</i>
Intervener length	Short (2 words)	1	117	0.008
	Long (7 words)	0	120	0.000
Intervener type	Objects	1	117	0.008
	Adverbials	0	120	0.000
Sentence type	Question	0	119	0.000
	Statement	1	118	0.008

As demonstrated in Tables 3.9 and 3.10, secondary *que* insertion is not an adopted strategy in either Cuban or Colombian groups. This finding is robust across all variables, intervener length, intervener type and sentence type. For Cuban Spanish, the proportion *que* doesn't exceed 0.021 across any of the conditions. For Mainland Spanish, the proportion *que* doesn't exceed 0.008 across any of the conditions. In Table 3.11, we highlight the similar between groups behavior.

Table 3.11. Sentence completion task, between groups comparison

Variables	Levels	Proportion <i>que</i> Cuban	Proportion <i>que</i> Colombian
Intervener length	Short (2 words)	0.005	0.008
	Long (7 words)	0.015	0.000
Intervener type	Objects	0.010	0.008
	Adverbials	0.010	0.000
Sentence type	Question	0.021	0.000
	Statement	0.000	0.008
	<i>Mean</i>	0.010	0.004

The Cuban and Colombian groups' mean proportion of items with *que* insertion is 0.010 and 0.004, respectively. This entails that 99% and 99.6% of items are produced without secondary *que*. No further analysis is required. We conclude that secondary *que* expression neither varies across dialect (RQ2) nor is it an adopted strategy to facilitate complement integration (RQ3).

3.5 DISCUSSION

The present chapter offered an experimental investigation on the linguistic phenomenon of recomplementation in Cuban and non-coastal Colombian Spanish. We adopted two offline production tasks: aural repetition and sentence completion. Together these tasks were assumed to serve as a window into whether secondary *que* expression is licensed by the grammar and whether it serves a facilitative function during complement integration. Three specific research questions motivated this chapter. The first question pertained to the overall grammaticality of overt C2. We predicted that secondary *que* expression was grammatical in Mainland Spanish. This expectation was supported by both experimental evidence (Frank, 2016; evidence from the previous chapter of this dissertation) and syntactic-theoretical accounts (e.g., Villa-García, 2015). The second research question scrutinized any divergence between groups that might be interpreted as dialectal variation. For (RQ2), we predicted that recomplementation was a locus of morphosyntactic variation between Cuban and Mainland Spanish. To arrive at this prediction, we weighed the expected results from (RQ1) against Frank and Toribio's (2017) conclusion that overt C2 is not grammatical in Cuban Spanish. The third and final research question considered whether overt C2 facilitated complement integration, independent of grammatical status. We predicted that overt C2 would facilitate complement integration,

which would be evidenced by C2 lexicalization in the null C2 condition of the repetition task and C2 insertion in the sentence completion task. We cited existing experimental (Casasanto & Sag, 2008; Echeverría & López Seoane, 2019) and theoretical (e.g., Ledgeway, 2000; Radford, 2018; Villa-García, 2019) literature arguing for an effect of intervener length on secondary *que* expression.

Results from the repetition task were robust. Among the many fixed effects and interactions defined by the model, only secondary *que* proved to be significant. Specifically, Cuban and Mainland Spanish participants repeat null C2 condition test items (94% and 99%) with greater accuracy than the overt C2 condition test items (26% and 23%). We conclude that overt C2 is not licensed by Spanish grammar, which does not support our prediction for RQ1. Further, a lack of effect for group and lack of interactions between group and all other fixed effects disconfirms our prediction for (RQ2). We find no evidence of recomplementation as a locus of dialectal variation. What's more, both the Cuban and Mainland Spanish groups repeated the short (60%, 53%) and long (62%, 67%) conditions without a significant difference in accuracy. This lack of intervener length effect, along with ceiling performance in the repetition accuracy of the null C2 condition, and an inverse relationship between age of participant and repetition accuracy of the overt C2 condition suggests that secondary *que* does not facilitate complement integration. This conclusion disconfirms our prediction for (RQ3). In sum, it can be said we struck out on our research question predictions (0/3). The supplemental sentence completion task only provides further confirmation. Namely, we find no evidence that secondary *que* is adopted as a production strategy. Specifically, for the Cuban and Mainland Spanish groups, 99% and 99.6% of the test items are produced without C2 insertion. This finding is robust across dialects and all other fixed effects, including intervener length.

The broader goals of this study were to address Frank and Toribio's limitations and to further scrutinize the gap between theoretical accounts and experimental findings. In terms of the former, we contributed a Colombian Spanish group to serve as a proxy for Mainland Spanish. We also offered a novel analysis of the data in order to account for linguistic microvariation. Specifically, we analyzed the effects of group, secondary *que*, intervener length, intervener type, sentence type and relevant interactions. Our results partially support Frank and Toribio's conclusions. On the one hand, their suggestion of recomplementation as a locus of variation is not supported by our analysis. On the other hand, a lack of evidence that secondary *que* expression facilitates complement integration is replicated here. However, this latter suggestion comes with a caveat. Namely, to more directly speak to the potential benefit of secondary *que* expression to processing costs and memory, an online methodology, as well as supporting cognitive measures (e.g., working memory span) should be adopted. This limitation will be accounted for in Chapter 4, where we attempt to replicate Casasanto and Sag's (2008) finding of an intervener length effect via the self-paced reading paradigm, with working memory span considered.

Our second aim was to leverage the findings of this chapter to speak to the divide between theoretical accounts and experimental findings. This gap refers to the tension between the case for grammaticality/optionality of secondary *que* expression on the one hand, and marginal acceptability or ungrammaticality on the other. The former argument is advanced by syntactic-theoretical accounts (e.g., Demonte & Fernández-Soriano, 2009; Villa-García, 2015) and the latter advanced by experimental findings (e.g., Casasanto & Sag, 2008; Frank, 2016). Unfortunately, a satisfying resolution remains elusive. Had we found evidence of dialectal variation, where C2 is licensed by the Colombian Spanish grammar but not the Cuban Spanish grammar, we could make the case that the gap is accounted for by the perceived Peninsular Spanish bias in the theoretical literature.

However, we found no such evidence. A second explanation could have been found in the lack of attention to potential microvariations within the phenomenon of recomplementation. Perhaps some conditions within intervener length, intervener type or sentence type are licensed by the grammar, while others are not. Further, these microvariations might differ between groups. Alas, we found no evidence of linguistic microvariation, neither between nor within groups.

A future study might operationalize methodological effects. Informal non-quantitative methods have come under recent scrutiny (see accuracy of judgment and researcher bias, e.g., Edelman & Christiansen 2003; Ferreira, 2005; Gibson & Fedorenko 2010a, 2010b). Formal experimental methods have not fared much better (e.g., see replication crisis in Cognitive Sciences; Aarts, 2015). Furthermore, studies that have compared data collected informally versus formally have found lower rates of convergence in Spanish (~75%) as compared to English (95%) (Ortega-Santos, 2020a; Sprouse et al., 2013). Thus, if we are to continue to investigate the issue of grammaticality in recomplementation, methodological effects seem a promising path forward. We also cannot ignore the possibility of task effect in this chapter. Specifically, the acceptability judgment task in Chapter 2 rated overt C2 within the range of marginal acceptability in Colombian Spanish (2.82 for questions and 2.99 for statements on a 1-totally acceptable to 7-totally unacceptable scale). This more nuanced take on grammaticality might suggest that an elicited imitation task, with a binary dependent measure interpreted as licensed or unlicensed by the grammar, is not sensitive enough to investigate recomplementation. Independent of data collection methods, we hope to have advanced the message that one must be wary of a Mainland or Peninsular bias in the literature. Theory must be informed by understudied dialects and microvariations must be scrutinized. In Chapter 4, we move

beyond the question of grammaticality and investigate whether secondary *que* reduces processing costs associated with complement integration during online comprehension.

Chapter 4: Incremental Sentence Processing Study

The previous chapter combats the Peninsular bias in the recomplementation literature by investigating Cuban and Colombian varieties. Overall results support the conclusion that overt C2 is not licensed by Colombian and Cuban Spanish grammars. Further, C2 lexicalization does not facilitate complement integration in an offline production task. Importantly, the possibility of task effect cannot be ruled out. In summary, by investigating understudied dialects, we found evidence that syntactic-theoretical assumptions of limited variation in recomplementation is justified. We must now consider the role of experimental task effect in promoting the divide in theoretical accounts and experimental findings. The present chapter adopts an online self-paced reading method. Further, it builds on the previous chapter by investigating complement integration, as well as the effect of individual differences in working memory capacity on reading times. Both memory-based and expectation-based psycholinguistic models of recomplementation are considered.

4.1 INTRODUCTION

The prediction of upcoming information plays a crucial role in syntactic parsing (e.g., see Kuperberg & Jaeger, 2016 for a review). However, what the prediction precisely entails and the relationship between forward-looking (e.g., anticipation) and backward-looking (e.g., retrieval) processes during real-time sentence processing continues to be a topic of debate. Parsing accounts that emphasize the forward-looking processes can be described as *expectation-based*; those that stress the backward-looking processes can be called *memory-based* (Staub, 2010).

Expectation-based accounts of syntactic parsing are grounded in a comprehender's real-world experience or their frequency of exposure to certain words and structures (e.g., Frazier, 1987; Frazier & Fodor, 1978; Hale, 2001; Levy, 2008; Mitchell et al., 1995). Processing complexity is related to the degree to which new information (mis)aligns with expectations. Memory-based accounts, on the other hand, are grounded in the comprehender's limited resource availability for information storage and retrieval (e.g., Gibson, 1998, 2000; Grodner & Gibson, 2005; Warren and Gibson, 2002). Processing complexity is then related to the strain on working memory, which is modulated by distance or linguistic complexity.

The present chapter investigates expectation-based versus memory-based accounts of recomplementation with new experimental evidence. The latter account is initially proposed in Casasanto and Sag (2008). As reviewed in Chapters 2 and 3, the authors investigate the online comprehension of multiple *that* constructions in English, with long and short intervening material, as in (4.1a) and (4.1b):

- (4.1) a. *Long intervener condition*
John reminded Mary **that** after he was finished with his meeting (**that**) his brother would be ready to leave.
- b. *Short intervener condition*
John reminded Mary **that** soon (**that**) his brother would be ready to leave.

Interpreting the results of a self-paced reading task, they find an interaction between length of material sandwiched between the complementizers (henceforth: intervener length) and the lexicalization of the secondary complementizer (C2). Specifically, reading times in the critical region *brother* (i.e., the embedded subject) were significantly faster in (4.1a) when C2 was overt as compared to null. The same effect was not found in (4.1b). They argue that secondary *that* facilitates integration of the complement by reiterating the information and function of the primary complementizer.

Casasanto and Sag frame their results in Distance Locality Theory, where online processing costs are a function of the distance between predicted dependencies held in memory and the integration of the dependent information (e.g., Gibson, 1998). In the case of recomplementation, the notion of “predicted dependencies” entails that the high complementizer (C1) predicts a complement (e.g., embedded subject or embedded verb in the case of pro-drop languages like Spanish). This dependency remains unresolved until complement integration, at which point old information must be retrieved. C2 lexicalization (as a reiteration of C1) reduces the length of the dependency or the retrieval distance to zero, lowering the strain on working memory prior to complement integration.

As will be elaborated later in this chapter, Casasanto and Sag’s (2008) analysis invites scrutiny on multiple levels, including their interpretation of the data and the assumption of secondary complementizer ungrammaticality. In response, we offer an alternative expectation-based account of recomplementation, where conventional patterns of C2 lexicalization can be probabilistically predicted along an intervener length constraint (e.g., Levy, 2008). We ground our proposal in the syntactic-theoretical accounts, experimental findings and written and spoken corpora analyses that converge on the conclusion that C2 lexicalization is correlated with intervener length (e.g., Casasanto & Sag, 2008; Echeverría & López Seoane, 2019; Ledgeway, 2000; Radford, 2018; Villa-García, 2019). Then we propose that a highly predictive parallel parser would probabilistically predict C2 lexicalization with greater certainty as intervener length gets longer. As in traditional expectation-based accounts, relative ease of integration is related to the degree to which the input aligns with expectations. Thus, when the expectation of C2 lexicalization is not met, complement integration is associated with greater processing costs. Conversely, reaction times are faster (i.e., lower) when events are predictable and expected. This chapter offers new experimental evidence, which informs our decision of

whether a memory-based or expectation-based model best accounts for recomplementation. As exploratory contributions, we also investigate whether matrix verb biases are used to anticipate new information and the relationship between individual differences in working memory capacity and performance. We begin with a selective review of the literature on syntactic parsing.

4.2 SYNTACTIC PARSING

There are several nuances to be considered in the relationship between forward-looking processes and backward-looking processes. Such topics include but are not limited to the ability of the parser to entertain one (serial) or multiple (parallel) syntactic structures at a time; the relative depth (i.e., syntactic algorithms) or shallowness (i.e., speedy heuristics) of the parse; the ability of the parser to anticipate information beyond syntactic structure; the nature of commitment and certainty *vis-à-vis* prediction; and the role of working memory. We will consider several of these topics in turn.

4.2.1 Expectation-Based and Memory-Based Accounts

As discussed earlier, the starting off point of expectation-based accounts of syntactic parsing is the comprehender's real-world experience. The assumption is that experience or amount of exposure to certain words and structures facilitates the creation of a database of frequency metrics that inform prediction (e.g., Frazier, 1987; Frazier & Fodor, 1978; Hale, 2001; Levy, 2008; MacDonald, 2013; Mitchell et al., 1995). Processing complexity is then related to the degree to which the input aligns with expectations. Quantifying experience for the purpose of empirical study is no easy feat. One criticism of expectation-based accounts is the use of frequency metrics measured from a corpus as a

proxy of experience. Roland et al. (2007) note that “differences in difficulties found in passives, actives, and the various cleft and relative clause constructions may not be explained by the individual frequencies of these constructions, but rather by the overall frequencies of higher level patterns such as subject verb object or agent verb patient” (379).

Frazier and Fodor (1978) proposed an expectation-based account that was “all or nothing”, where the parser entertains one syntactic structure at a time. They argued that in order to efficiently analyze new input as it presents itself incrementally, the parser has to follow fairly simple heuristics. According to the minimal attachment principle, comprehenders initially organize words using a single structure with the fewest number of syntactic nodes possible. When a rudimentary structural outline no longer fits new evidence, the parser creates a new, more complex structure. This process of reanalysis repeats itself in a serial fashion (also see van Gompel, Pickering, & Traxler, 2001; Traxler, Pickering, & Clifton, 1998).

Other expectation-based accounts propose that multiple structures can be hypothesized at any given moment (e.g., Hale, 2001; Jurafsky, 1996; Levy, 2008; Lewis, 2000). According to Kuppenberg and Jaeger (2016), one major difference between parallel and serial expectation-based accounts is that in the former, reanalysis is akin to a redistribution or reweighting of anticipated structures. In the latter, reanalysis is akin to moving on to the next option. One example of a parallel account is Levy’s (2008) probabilistic framework. In this framework, each of the predicted hypotheses have different degrees of probability and are ordered along a probability distribution (see Bayesian Surprise; Doya et al., 2007). As new information presents itself, probability distributions change. This change could entail a new ordering of hypotheses and/or a stronger commitment to an existing ordering. Importantly, the extent of the shift is correlated with processing complexity. The probability distribution is updated after every new piece of

information, a cycle that repeats itself as the parser moves from relative uncertainty at the beginning of the sentence to relative certainty by the end.

The specific role of semantic information in expectation-based accounts has been a topic of debate. Some have adopted a “structure first” approach, where verb biases are only used as a revision mechanism (e.g., Ferreira & Henderson, 1990; Frazier & Rayner, 1982). Others argue that verb biases are essential to anticipatory processes (e.g., Altmann & Kamide, 1999, 2007; Kamide et al., 2003). Altmann and Kamide (1999) tracked eye movements as participants looked at several visual scenes and listened to sentences like (4.2a) and (4.2b):

- (4.2) a. The boy will eat the cake.
b. The boy will move the cake.

The scene that accompanied the sentences above was of a boy surrounded by several objects, only one of which was edible, i.e., the cake. The authors found that saccades to the image of the cake were initiated earlier when the main verb was “eat” as opposed to “move”. They conclude that sentence processing is highly predictive in nature and that even at the earliest stages of comprehension, there is evidence of an interaction between syntax, verb meanings, and real-world contexts (also see Ford et al., 1982; Gibson & Wu, 2013; Garnsey et al., 1997; Wilson & Garnsey, 2009).

Memory-based accounts, on the other hand, are grounded in a comprehender’s limited memory resource availability for information storage and retrieval (e.g., Gibson, 1998, 2000; Grodner & Gibson, 2005; Warren and Gibson, 2002). The source of processing complexity is directly related to the strain on working memory. As introduced in the beginning of this chapter, the Dependency Locality Theory (henceforth, DLT) is a well-known example of a memory-based account of processing complexity (e.g., Gibson, 1998). DLT specifically proposes that processing complexity is a function of two distance-related

dimensions: (1) the storage of the predicted syntactic/semantic content and the incomplete dependencies; and (2) the manipulation or attachment of new material that is dependent on stored material. In short, processing costs are a function of the distance between predicted dependencies held in memory and the integration of the outstanding dependent information. Crucially, the longer the distance of the dependency, the greater the strain on working memory and the higher the integration cost. Gibson and colleagues specifically argue that the integration cost $I(n)$ associated with two elements engaged in a syntactic dependency is equal to the distance n between them, $I(n)=n$. They offer that n can be measured by counting letters, syllables, words or constituents.

Importantly, expectation-based and memory-based accounts need not be applied mutually exclusively to a linguistic phenomenon. Staub (2010) explicitly tested a memory-based account (e.g., Gibson, 1998) against an experience-based expectations account (e.g., Levy, 2008). To investigate this question, he considered whether the processing difficulty in reading object relative clauses (as opposed to subject relative clauses) appears on the verb or the subject of the relative clause, as represented by example (4.3ab) below:

- (4.3) a. *Subject Relative Clause (SRC) condition*
The reporter that __ attacked the senator admitted the error.
- b. *Object Relative Clause (ORC) condition*
The reporter that the senator attacked __ admitted the error.

In the object relative clause condition (ORC), *the reporter* is the theme/patient of the verb attacked. It is thus linked to a phonologically empty element in the object position. The author claims that under a memory-based account, the reporter must be held in working memory until the verb attacked is encountered, at which point it must be retrieved. Thus, the ORC is more difficult to process than the SRC because of the increased retrieval distance, or the length of the intervening material (i.e., the senator). Under an expectation-

based account, (4.3b) is more difficult to process than (4.3a) because it is less common than the latter (Roland, Dick, & Elman, 2007). Specifically, after encountering the relativizer *that*, an SRC is preferentially expected over an ORC. These two accounts offer unique predictions for the locus of the processing difficulty. The memory-based prediction suggests that the onset of difficulty in the ORC is at *attacked*, at which point the reporter must be retrieved. The experience-based prediction suggests that the onset of difficulty is localized at the unexpected constituent *the senator*, at which point the construction can no longer be the more frequent SRC construction. The author finds evidence that both regions are difficult to process and concludes that both experience-based expectations and processes of memory retrieval contribute to reading comprehension difficulty of object relative clauses. In other words, being surprised during word-by-word processing and the retrieval and integration of a long-distance dependency incurs a substantial processing cost.

4.2.2 Committing to a Prediction

The role of distance is front and center in memory-based accounts-- dependency distance equals processing complexity. The role of distance is less straightforward though no less important in expectation-based accounts. One outcome of Levy's (2008) framework is an inherent time/strength relationship. For example, if you incorrectly increase your certainty of a hypothesis with each new piece of information (i.e., you strengthen your prediction error), integration will be increasingly costly at the moment of disambiguation. The correlation between time committed to an analysis and strength of commitment has been adopted as a possible explanation for processing complexity across a number of linguistic phenomena, including garden path sentences (e.g., Bailey & Ferreira, 2003;

Christianson et al., 2001; Ferreira & Henderson, 1991). See examples (4.4a-c), where the intervener length has been manipulated (e.g., Bailey & Ferreira, 2003):

- (4.4) a. When the gardener bathes his poodle joins him.
c. When the gardener bathes his poodle with the soft fur joins him.
c. When the gardener bathes his poodle uh uh joins him.

In (4.4b) and (4.4c), extra material has been included between the head of the misanalysed phrase *poodle* and the disambiguating word *joins*. If the commitment to the wrong interpretation (i.e., the gardener is bathing his poodle) is strengthened in these manipulations, then there should be a greater integration cost at the disambiguating region. After either reading (Ferreira & Henderson, 1991) or listening to (Bailey & Ferreira, 2003) a series of sentences, participants were asked to judge them as “grammatical” or “ungrammatical”. Results from these two studies support the Time/Strength Hypothesis, where sentences with extra material were more often deemed ungrammatical. Furthermore, the position of the extra material was a significant factor. For example, when the disfluency *uh uh* was moved to the position directly in front of *poodle*, the degradation in grammaticality judgment disappeared. This can be explained by the fact that the misanalysed phrase and the disambiguating region (i.e., *poodle* and *join*) are now adjacent. The authors further conclude that interruptions of many types influence the parser, where the phrase that immediately follows is more likely to be interpreted as a new clause.

The phenomenon of subconsciously making a strong commitment to a prediction prior to the availability of new information is called *predictive preactivation*. According to Kuperberg and Jaeger (2016), this is when we use a “...high-level event hypothesis to predictively pre-activate [a representation] prior to the bottom-up input reaching [said representation]” (39). Predictive preactivation affords the parser an early start on

processing, thus facilitating eventual integration of new information. On the other hand, the authors argue that a strong and early high-level (e.g., real-world knowledge, frequent structures) commitment alongside relative uncertainty of incoming information can also lead to misinterpretations (see Good Enough processing, e.g., Ferreira & Patson, 2007). Ferreira's (2003) seminal work found that highly educated native speakers of English often misinterpret non-canonical sentences, such as passive constructions like (4.5).

(4.5) The dog was bitten by the man.

Specifically, participants interpreted (4.5) to mean *the dog bit the man*. She argues that the common knowledge that dogs bite men, along with the misleading surface word order cues, trigger a noun-verb-noun (NVN) parse, which further triggers a 1st-NP-is-the-agent heuristic. The author concludes that the comprehension system uses syntactic algorithms and speedy heuristics when processing a sentence. Evidence for Good Enough processing has also been derived from the disruption of global sentence interpretation by local meaning in garden pathing (e.g., *While Mary bathed the baby played in the crib*); the use of a probability heuristic to dispel ambiguity (e.g., *The singer saw the audience member with the binoculars*); and the comprehender's tendency to interpret intended meaning rather than the actual meaning (e.g., *Where should the authorities bury the survivors*, Barton & Sanford, 1993)

In summary, the Good Enough parser provides a less detailed representation based on lexical-semantic information, canonical word orders and other surface cues (e.g., Ferreira, 2003; Ferreira & Patson, 2007; Ferreira et al., 2002; Sanford & Sturt, 2002). Ferreira and colleagues question the underlying assumption that the goal of the language comprehension system is to deliver an accurate and detailed representation of a speaker's utterance. They offer that during communication people may apply superficial heuristics. This shallow parser is associated with a speed-accuracy tradeoff, where reliance on speedy

and economical heuristics can be at the expense of an accurate interpretation. Ferreira (2003) argues in favor of a dual processing mechanism, where the output of the full parser and the good enough parser work in tandem. Either "... [they] operate in parallel, ...the heuristics are used only when the algorithms become hard to apply, ...[or] algorithms are used only when the comprehender lacks confidence in the interpretation created from the heuristics" (169). The precise nature of the potential interaction between good enough and full syntactic parsing in a dual processing mechanism is beyond the scope of this chapter. We do however note that strong predictions not only impact the integration cost of new information but also overall comprehension accuracy.

4.2.3 Working Memory

Lastly, we review the fundamental role that working memory plays in both forward-looking and backward-looking processes. According to Marull (2017a), "...working memory is the space in which information from multiple sources such as morphology, syntax, semantics, and discourse information are retrieved from lexical entries and integrated to create structure". She continues, "... the capacity for any individual to actively consult multiple information types in parallel at each stage of sentence processing is fundamentally linked to his or her working memory capacity" (51). Baddeley and Hitch (1974) proposed a model of working memory. According to them, between sensory memory and long-term memory there is a space for (1) temporary storage and (2) retrieval and manipulation of information. Their model for the short-term storage and manipulation of information originally included three parts: the central executive, or the decision maker, which then draws on the visuospatial sketchpad and the phonological loop. The visuospatial sketchpad refers to the creation of mental imagery, while the phonological

loop refers to a mental repetition or rehearsal. Later, Baddeley (2000) added a fourth component to the model, the episodic buffer, which facilitates the integration of all the parts of the model into a coherent narrative or episode (also see Baddeley, 2003, 2007).

The effect of individual differences in working memory capacity on syntactic parsing is well documented (e.g., Daneman and Carpenter, 1980; Huettig & Janse, 2016; Janse & Jesse, 2014; Just & Carpenter, 1992; Macdonald et al., 1992; Rönnerberg et al., 2013). In the seminal work of Daneman and Carpenter (1980), the authors find evidence in favor of a correlation between working memory span and language comprehension. They posit that participants with lower reading comprehension scores are able to recall fewer words in a sentence recall task because reading inefficiencies limit the amount of information they can hold in working memory. Just and Carpenter (1992) propose a Capacity Theory of working memory, where limited resources drive a constant tradeoff between storage and processing allocation. When resource demands exceed available supply, resources allocated to old information held in storage may be reallocated to facilitate expectation and manipulation of new information. One outcome of the Capacity Theory is that comprehenders with more available resources are able to hold more parallel activations in any given moment and make better predictions. In support, MacDonald et al. (1992) finds that comprehenders with higher working memory capacity are able to maintain parallel structure activation in temporarily syntactically ambiguous sentences for a longer period of time (e.g., *The soldiers warned about the dangers conducted the midnight raid*). Thus, they are able to integrate less frequent resolutions with greater ease, when compared to comprehenders with lower working memory capacity who inactivate the less frequent resolution more quickly.

More recently, Huettig and Janse (2016) investigated 105 participants' ability to use gender marking to anticipate target objects via visual world paradigm. For example, as

participants listened to the instructions, *Look at the displayed piano* (recited in Dutch), they saw four objects on the screen. The target object was a piano and the other three were distractors unrelated to the prompt. Crucially, gender information from the definite determiner could be used to predict the target object only. They found that those with higher working memory capacity displayed significantly more target anticipatory eye movements.

Janse and Jesse (2014) investigated the ability of older populations (range 64 to 89 years) to use contextual information to facilitate real-time comprehension of spoken words. A phoneme monitoring task asked participants to monitor recorded speech for target phonemes (e.g., /p/ in *pill*, as in the sentence, *The circus artist had a **pill** for years that kept her nerves under control*) by pressing a key as quickly as possible when the phoneme was detected. Importantly, the contextual probability of the target words was operationalized as a continuous variable. The authors found that the ability to use context information in the timely recognition of target words was modulated by working memory scores on a reverse digit span task. They interpret this finding as evidence that an increased ability to store and manipulate semantic representations supports efficient processing.

Rönnberg et al. (2013) expands on the Ease of Language Understanding model, which provides evidence of the relationship between working memory capacity and early attention mechanisms, hearing impairment, accommodation of different task demands (e.g., noise-in-speech comprehension), and general effort as measured by pupil dilation in high-demand listening environments. In summary, these studies, and many more, point to the central role that working memory plays in online sentence comprehension.

4.3 COMPETING ACCOUNTS OF RECOMPLEMENTATION

4.3.1 Memory-Based Accounts

As mentioned in Section 4.1, Casasanto and Sag (2008) proposed a memory-based account of recomplementation. Specifically, they frame their results in the Distance Locality Theory (e.g., Gibson, 1998), where the high complementizer (C1) predicts a complement (e.g., embedded subject or embedded verb), which remains unresolved until complement integration. Importantly, C2 lexicalization lowers the strain on working memory and the processing complexity associated with complement integration by reducing the dependency distance to zero (see examples [4.1a] and [4.1b]). The authors also claim that C2 lexicalization is ungrammatical and that there is a grammar/processing tradeoff with an *advantage of the ungrammatical* in the long condition (compare [4.1a] to [4.1b]). They argue that framing the online comprehension of recomplementation along these parameters accurately explains their results, as demonstrated in Figure 4.1.

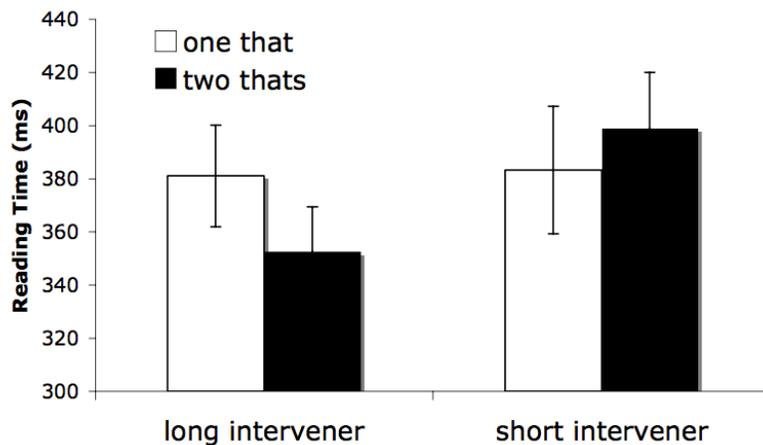


Figure 4.1. Casasanto and Sag's (2008) self-paced reading results

They find a significant interaction between intervener length and C2 lexicalization factors. Post hoc analysis shows that reading times in the critical region (i.e., the embedded

subject) were significantly faster in the overt C2 long condition when compared to the null variety. The same effect was not found in short condition. In fact, for the short intervener condition, the authors find an (insignificant) trend in the opposite direction, where overt C2 is more difficult to process than the null variety.

Casasanto and Sag’s analysis merits closer inspection. First let us consider that the strict interpretation of Dependency Locality Theory (DLT) versus DLT plus the grammar/processing tradeoff hypothesis (henceforth: DLT+) make vastly different predictions, as highlighted in comparing Figures 4.2 and 4.3.

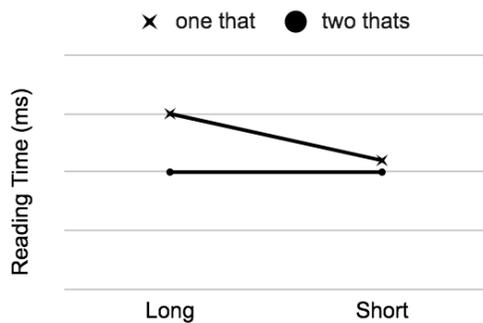


Figure 4.2. Interaction predicted by DLT

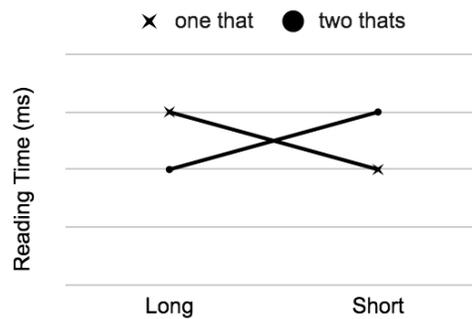


Figure 4.3. Interaction predicted by DLT+ (Casasanto & Sag, 2008)

DLT predicts that the overt C2 (or two thats) conditions would be read with equivalent ease. Specifically, when C2 is lexicalized, the dependency distance between complement and complementizer is reduced to zero. This is entirely independent of intervener length. DLT further predicts that the null/short condition would be read faster than the null/long one. In the absence of overt C2, processing complexity increases as a function of distance. DLT+, on the other hand, makes different predictions. As the author’s note, overt C2 is not licensed by the grammar. Crucially, following the tenets of DLT, C2 lexicalization facilitates complement integration by reducing the dependency distance to zero. Casasanto and Sag argue that when dependency distance is long, the C2 lexicalization

benefits (i.e., reduction in processing complexity) outweigh the costs (i.e., anomalous/ungrammatical C2). This processing/grammar tradeoff is depicted in Figure 4.2 by the faster reading time in the overt C2 long as opposed to short condition. DLT+ also predicts faster reading times in the null C2 short as opposed to the long condition. Given the absence of a grammatical violation, this prediction is identical to the prediction made by DLT.

However, the shape of neither of these interactions fits Casasanto and Sag's findings particularly well. Unfortunately, the authors only report the significant interaction and the results of the corresponding post hoc analysis as discussed above. However, after taking both their available statistical report and their graph (see Figure 4.1) into account, we deduce the following interaction:

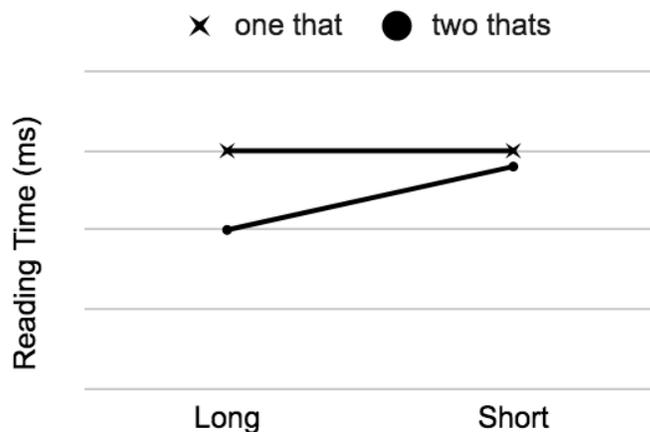


Figure 4.4. Alternative interpretation of the interaction

First, we assume no significant main effect for intervener length or C2 lexicalization. No such data is reported and their graph generally supports this assumption. Second, as clearly reported in their post hoc analysis, overt C2 items are read significantly faster than null C2 items. No significant effect was found in the short condition. We disregard the reported trend in the opposite direction because it does not reach significance.

Further, if we began to discuss insignificant trends, then we would argue that the overall trend in the data for the short condition to be more complex than the long condition is difficult to reconcile with DLT, even in an ungrammatical C2 environment. Further, the seemingly equivalent complexity between null/short and null/long conditions is nearly irreconcilable. Yet, these two data points are left unmentioned in their analysis and discussion.

In addition to critiquing the interpretation of the data, we challenge Casasanto and Sag's claim of an *advantage of the ungrammatical* or a grammar/processing tradeoff, which assumes overt C2 is a grammatical violation. This assumption is in fact important for their analysis and accounts for the difference in shape between the DLT and DLT+ interaction. But this assumption is unfounded and highly controversial. They cite as evidence only a decrement in acceptability judgment found in a supplemental task. Moreover, they provide no formal support for the claim, simply stating that the second "...complementizer is not an option made available by any principled grammar (formal, pedagogical, etc.) of English, and in fact it does not provide any new grammatical information - it simply reiterates information provided by the original complementizer" (602). As reviewed in Chapters 2 and 3, there is an extensive body of syntactic-theoretical work that argues for the grammaticality of the secondary complementizer (e.g., see Villa-García, 2019 for a unified Spanish and English account; also, Radford, 2018). What's more, leading grammatical accounts of the secondary complementizer argue that it is in a specifier/head relationship with Topic. It certainly provides *new* grammatical information and is far more than a reiteration of the first, though it can also serve that function (e.g., Radford, 2013; Villa-García, 2015; also see the polyvalent complementizer, Villa-García, 2019).

4.3.2 An Expectation-Based Account

Critiques on Casasanto and Sag's data interpretation and underlying assumptions motivate us to propose an alternate model of recomplementation, one that better accounts for the shape of the interaction as we understand it (see Figure 4.4) and one that does not assume a grammatical violation and does not commit to a singular C2 function. Before proceeding to the proposed account, it proves useful to review a relevant subset of the recomplementation literature. Syntactic-theoretical accounts, experimental findings and written and spoken corpora analyses converge on the conclusion that C2 lexicalization is correlated with intervener length (e.g., Casasanto & Sag, 2008; Echeverría & López Seoane, 2019; Ledgeway, 2000; Radford, 2018; Villa-García, 2019). Radford's (2018) English corpus displays on average 5.9 words per phrase that is flanked by a complementizer. Echeverría and López Seoane (2019) find 124 instances of recomplementation in the Spanish text *El conde Lucanor*. Of the instances with short intervener length (i.e., 1-3 words), only 6/61 (10%) display overt C2. Of the instances with long intervener length (i.e., 4 or more words), 43/63 (68%) display overt C2. Ledgeway's (2000) Moving Complementizers formal account notes that secondary *que* is more likely to be pronounced when the dislocated material is heavy. Casasanto and Sag (2008) provide experimental evidence of an interaction between length and C2 lexicalization, where C2 lexicalization facilitates processing when the intervener is long.

This literature review makes a compelling case for what we will define as the **C2 complexity correlate**, where conventional patterns of C2 lexicalization can be probabilistically predicted along a length-of-intervener constraint. Thus, we propose an expectation-based account of recomplementation. As discussed earlier, the starting off point of expectation-based accounts of syntactic parsing is the comprehender's real-world experience. We refer to the convergence in the literature on C2 as a complexity correlate

and assume that as intervener length increases, the probability of hearing overt C2 also increases. We propose that a predictive parallel parser would activate multiple hypotheses with different degrees of probability along a frequency distribution. The parser would probabilistically predict C2 lexicalization with greater certainty as intervener length increases. In other words, commitment to the overt C2 representation strengthens with each passing word. Given that processing complexity is equal to the shift in probability distribution, we hypothesize that as intervener length increases so does the complement integration cost when the expectation of C2 lexicalization is not met. Likewise, integration cost is lower when expectations are met.

In review, memory-based (i.e., DLT or DLT+) and expectation-based accounts of recomplementation posit that processing complexity increases as a function of distance. The former adopts a backward-looking “dependency retrieval” account, while the latter adopts a forward-looking “unmet expectations” account. Conversely, they both predict that the benefit of C2 lexicalization to the parser increases as a function of distance. In the memory-based accounts, C2 lexicalization reduces complexity by erasing the distance dependency. In the expectation-based account, C2 lexicalization reduces complexity by “meeting expectation” of a complexity correlate. If the literature is correct and C2 can be probabilistically predicted as a function of distance, then the burden is on the detractor to explain why a parser that predicts oncoming information wouldn’t take such information into account. Crucially, the expectation-based account is consistent with the syntactic-theoretical recomplementation literature that convincingly argues for the grammaticality and multifunctional nature of C2 (e.g., Radford, 2013, 2018; Villa-García, 2015, 2019). Memory-based accounts, on the other hand, are not. DLT+ posits the ungrammaticality of C2 lexicalization, as well as its singular function to reiterate C1.

More generally, the benefits of investigating recomplementation from a forward-looking perspective are many. For example, we can consider probability distributions earlier on in the sentence. Recomplementation as a left-dislocated phenomenon represents an infrequent, non-canonical form. Thus, it might take time for the probability distribution to reweight appropriately from a more frequent hypothesis (single complementizer) to a less frequent one (double complementizer). As a result, this shift in probability distribution in the short intervener condition be as costly as an unmet C2 expectation in the long condition. Additional benefits of the expectation-based account include the exploration of the effect of semantic information on hypothesis formation and probability distribution. While earlier processes in the incremental processing of recomplementation are not the primary focus of the present chapter, we do investigate recomplementation questions and operationalize matrix verb type (*preguntar* ‘to ask/wonder’ versus *decir* ‘to say/tell’) for the very purpose of exploring this variable.

4.3.3 Research Questions

In the previous section, we reviewed competing proposals. We first highlighted the shortcomings of Casasanto and Sag’s (2008) memory-based account. Then we proposed an expectation-based model and offered general insights into the advantage of forward-looking accounts. In the following sections we offer experimental evidence to empirically inform our decision of whether a memory-based or expectation-based model best accounts for recomplementation. Below, we offer the research questions and hypotheses that will drive the remainder of the chapter.

- (RQ1) Are matrix verb biases used to anticipate new information?
- (RQ2) Do memory-based or expectation-based models better account for the shape of the interaction between intervener length and C2 lexicalization?

(RQ3) What is the relationship between individual differences in working memory capacity and online performance?

For (RQ1), we offer a preliminary investigation into whether the onset of the left-dislocated material is more difficult to integrate after the verb *preguntar* ‘to ask’ or *decir* ‘to say/tell’ in recomplementation questions. We hypothesize that if verb biases are used in an expectation-based account to anticipate new information or predict structure, then processing complexity will be greater after the verb *preguntar* than *decir*. In other words, a comprehender would be more surprised by *Ella me **pregunta** que esa receta* than by *Ella me **dice** que esa receta*, see (4.6) and (4.7), respectively.

(4.6) Preguntar que -- prediction: [+wh] -- esa receta ×
To ask that that recipe

(4.7) Decir que -- prediction: [-wh] -- esa receta ✓
To say that that recipe

Thus, we would predict an integration penalty in the critical region *esa* (n), which could spill over into *receta* (n+1). While the verb *decir* ‘to say/tell’ can select for an indirect question, it typically introduces a *that*-clause statement (e.g., Demonte & Fernández-Soriano, 2009; Plann, 1982; Suñer, 1992, 1993; Rivero, 1980, 1994). On the other hand, the verb *preguntar* ‘to ask’ obligatorily selects for indirect questions, where left-dislocation is the non-canonical, less frequent form. If matrix verb information is not used to predict forthcoming information, then we do not expect to see an effect of verb type.

(RQ2) pertains to our primary area of investigation. We ask whether memory-based or expectation-based models of recomplementation better account for the shape of the interaction between intervener length and C2 lexicalization. We hypothesize that if the memory-based DLT account of recomplementation is accurate, then overt C2 conditions (both long and short) would be read with equivalent ease. Specifically, when C2 is

lexicalized, the dependency distance between complement and complementizer is reduced to zero. With regard to the null condition, we predict the short variety will be read faster than the long one. That is to say, in the absence of overt C2, processing complexity increases as a function of distance (see Figure 4.2). On the other hand, if DLT+ best accounts for the data, then the overt C2 condition will be read faster in the long as opposed to the short condition. Conversely, the null C2 condition will be read faster in the short as opposed to the long condition. This is the outcome of the processing/grammar tradeoff as applied to DLT by Casasanto and Sag (see Figure 4.3).

Lastly, if an expectation-based account of recomplementation is accurate, then the parser would probabilistically predict C2 lexicalization with greater certainty as intervener length increases. As a result, we hypothesize that when the expectation of C2 lexicalization is not met, complement integration costs increase as a function of distance. Similarly, when expectations are met, there will be a facilitation effect. Further, the short condition will be more costly in general. Recomplementation as a left-dislocated phenomenon represents an infrequent, non-canonical form. Thus, it takes time for the probability distribution to reweight appropriately from a more frequent hypothesis (single complementizer) to a less frequent one (double complementizer). This shift in probability distribution in the short condition may in fact be as costly as the unmet C2 expectation in the long condition (see Figure 4.4).

For (RQ3), we offer an exploratory look into the relationship between individual differences in working memory capacity and the online comprehension of recomplementation. Further, we use these results to test whether a memory-based or expectation-based account best fits the data. We hypothesize that if DLT or DLT+ best account for the data, then comprehenders with relatively lower working memory span would benefit most from the overt C2, particularly when the dependency distance is long.

In other words, the strain on available resources would be greater in that group. This is the scarcity of resources hypothesis. Contrastingly, if an expectation-based account of recomplementation is more accurate, then the comprehender with higher working memory capacity would be able to allocate more resources to anticipatory processes, such as the C2 complexity correlate. As a result, they would probabilistically predict C2 lexicalization as a function of length. When expectations are met, reading times would be faster. When expectations are not met, reading times would be slower. This is the more-resources-to-allocate-to-prediction hypothesis. Crucially, memory-based and expectation-based accounts make different predictions about the relationship between individual differences in working memory capacity and the benefit of C2 lexicalization.

4.4 THE STUDY

4.4.1 Participants

A total of 24 participants were recruited in Havana, Cuba in order to shed light on the aforementioned research questions and hypotheses. Selected participant metadata solicited from a language background questionnaire (see Appendix A) are summarized in Table 4.1:

Table 4.1. Selected participant metadata

Selected Information	Cuban Spanish (n=24)
Birthplace	Cuba
Current Residence	Havana
First language	Spanish
Parent's First Language	Spanish
Gender	11 male, 13 female
Age at testing	18-74 ($M= 33$)
Level of Education	18 university 4 technical school 2 high school
DELE Reading proficiency	$M= 45/50$ ($SD=3.4$, range: 35-48)
Working memory span	$M= 31.8$ ($SD = 12.4$, range: 14–52)

The participant group reported Spanish as both their first language and the first language of their parents. Further, the vast majority had a post-secondary school level of education. At time of testing, they held such occupations as student, teacher, engineer, accountant, journalist, physical therapist, army employee and retiree. In addition to a language background questionnaire, all participants completed an adapted version of the DELE (Diploma of Spanish as a Foreign Language) proficiency exam composed of a cloze passage and a multiple-choice vocabulary section from an MLA placement test, as in Appendix B (e.g., Montrul & Slabakova, 2003). Proficiency was elicited via the written mode to match the mode of the experimental self-paced reading task. Participants on average answered 45 out of 50 (90%) questions correctly, $SD = 3.4$, range: 35-48. Results suggest high proficiency in the written mode, with only three participants scoring below a 40.

In order to determine individual differences in working memory capacity we administered a visual reverse digit span task. Forward digit span tasks, where participants

store and recall digit sequences, are typically used to measure short-term memory. A reverse digit span task, where participants store, manipulate and recall sequences in the reverse order, is thought to be a more robust method for measuring working memory span (e.g., Baddeley, 2006). We adopted a visual mode, as opposed to an aural or bimodal medium, to match the exclusively visual nature of the experimental task. The visual reverse digit span task was administered on a laptop using the program PsychoPy (Pierce et al., 2019). This program controlled for the rate of stimuli presentation and automatically recorded and saved responses. The test began with the participant focusing their attention on a fixation point (+). After pressing the *enter* key, they saw a sequence of digits presented at a rate of one digit per second. When the screen displayed *answer now*, they responded verbally with the number sequence in the reverse order. The digit sequences began with a length of two digits. Each sequence length was repeated twice— with digits in a randomized order each time— before moving on to a longer length. This process continued until both trials within a given sequence length were recalled incorrectly. The maximum number of trials was 16 (2 sequences per digit lengths 2 through 9) and the total possible score was 88. Sample stimuli for a full list of trials are offered in Appendix G. We manually scored the results by adding the number of digits in the correct serial position for each sequence. For example, if the sequence was 34567 and the recall was 34657, then the participant would receive 3 out of a possible 5 points. The average score from 24 participants was $M = 31.8$ correct digits recalled, $SD = 12.4$, range: 14-52.

4.4.2 Experimental Methodology

The experiment itself consisted of a self-paced reading (SPR) task. According to Jegerski (2014), “nearly forty years after its development, SPR is still the most fundamental

experimental measure employed by psycholinguists interested in processing at or above the level of the sentence” (1). This method has been used to measure processing behaviors that pertain to ambiguity resolution within a sentence (e.g., garden pathing), non-canonical structures (e.g., passive constructions), ungrammatical forms (e.g., person/number agreement) and long-distance dependencies (e.g., relative clause attachment, recomplementation). The author goes on to explain that the underlying assumption of this method is that “...relatively longer reading times are taken as indications of processing difficulty, while faster reading times are interpreted as a sign that facilitation occurred” (4). The display method of the SPR task is typically noncumulative with word by word segmentation. In other words, there is a “moving window” effect with each passing word that is triggered by participant keyboard presses. This effect is demonstrated in Figure 4.5:

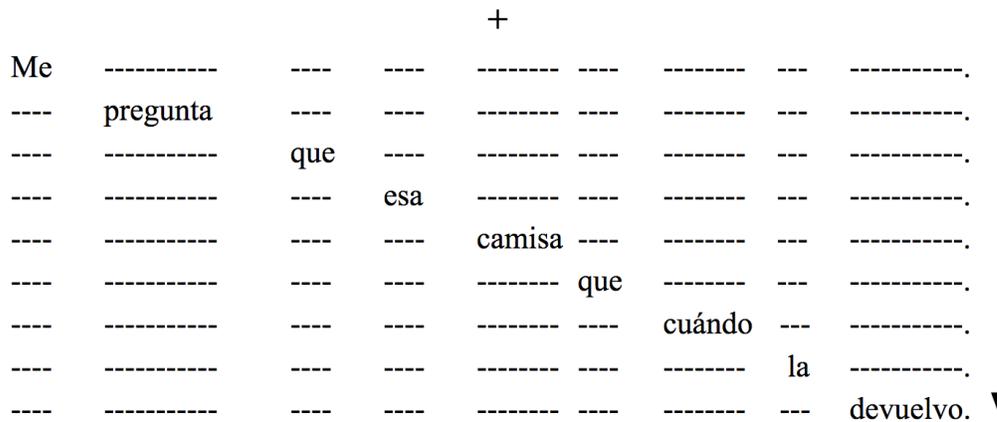


Figure 4.5. Moving window paradigm with word-by-word segmentation

Participants are first instructed to direct their attention to the fixation point (+), which (re)focuses their gaze at the center of the computer screen prior to each sentence. Each time they press the spacebar the next word will appear and the previous word will disappear. Although reading this way may seem strange, they are told to read the sentences in the most natural way possible, as if they were reading a newspaper or a magazine, and

to try to comprehend each sentence as fast as possible. Unbeknownst to the participant, the time between button presses is recorded. In order to get comfortable with this method, participants perform a short practice block of sentences. In our experiment, we have included six practice stimuli, targeting unrelated linguistic phenomena (i.e., adverb and negative particle placement and agreement morphology). As is typical in SPR experiments, stimuli are immediately followed by yes/no comprehension questions to ensure attention to task. All items where the comprehension question is answered incorrectly are not included in the final analysis and participants who score below a 70% overall are removed from the analysis altogether (e.g., Tucker et al., 2015). The experiment was designed and presented on the researcher's laptop using E-Prime 2.0 software (Psychology Software Tools, Pittsburgh, PA).

The design of our test items was inspired by Casasanto and Sag (2008) who investigated the benefit of the secondary complementizer in English recomplementation statements, manipulating intervener length and C2 lexicalization variables. Given that recomplementation literature, both syntactic-theoretical and functional is unified across Spanish and English (e.g., Villa-García, 2019), we are not investigating differences between languages. Rather, we examine Spanish because it facilitates an additional analysis of the effect of matrix verb bias on the prediction of new information. Namely, because non-ask/wonder verbs can introduce questions in Spanish, we are able to investigate both dislocated topic and complement integration within a single experiment. In summary, the three primary points of departure in our investigation when compared Casasanto and Sag (2008) are language, the sentence type and the additional matrix verb variable. Our three experimental variables resulted in 8 conditions. Six stimuli were created for each condition for a total of 48 master stimuli (see Appendix H). An example test item for each condition is represented in Table 4.2:

Table 4.2. Eight experimental conditions

Condition	Matrix Verb	C1	Intervener Length	C2	Sample Sentence
1	Decir	Overt	Short (2 words)	Overt	Él me dice que esa cocina que cuándo la reparo para mi comunidad. 'He asks me when I will repair the kitchen for the community.'
2	Decir	Overt	Short (2 words)	Null	Ella me dice que al secretario cuándo le ofrezco un descanso merecido. 'She asks me when I will offer a well deserved break to the secretary.'
3	Decir	Overt	Long (5 words)	Overt	Él me dice que al caballo de la carrera que cuándo le sirvo la comida nutritiva. 'He asks me when I will serve the race horse the nutritious food.'
4	Decir	Overt	Long (5 words)	Null	Él me dice que ese edificio con el mercado dónde lo construyo en el centro. 'He asks me where downtown I will construct the building with the market.'
5	Preguntar	Overt	Short (2 words)	Overt	Él me pregunta que al experto que cuándo le muestro mi trabajo cumplido. 'He asks me when I will show the expert my finished work.'
6	Preguntar	Overt	Short (2 words)	Null	Él me pregunta que esa película cuándo la muestro en mi fiesta. 'He asks me when I will show the movie at my party.'
7	Preguntar	Overt	Long (5 words)	Overt	Ella me pregunta que ese ensayo sobre mi historia que cuándo lo publico en una revista. 'She asks when I will publish the essay about my history in the magazine.'
8	Preguntar	Overt	Long (5 words)	Null	Ella me pregunta que esa casita de mi abuelo cuándo la pinto para mi familia. 'She asks me when I will paint my grandfather's house for my family.'

As demonstrated in Table 4.3, the matrix sentence, or introductory section, of the test items is composed of the following words: the first word of every test item is either the

pronoun *él* ‘he’ or *ella* ‘she’; the second word is the indirect object pronoun *me* ‘me’; the matrix verb is either the ask/wonder verb of saying *preguntar* ‘to ask’ or the non-ask/wonder verb of saying *decir* ‘to say/tell’; and the fourth word is the primary complementizer (C1) *que* ‘that’. In review, the introductory “matrix verb” section of each test item is maximally comparable with the exception of the test variable, matrix verb type. The two sections that follow are of particular interest because they represent critical regions for measurement of processing difficulty. All three sections are summarized in Table 4.3:

Table 4.3. Three sections for each test item

Section 1	Section 2	Section 3
<i>matrix</i>	<i>Topic (C2)</i>	<i>wh-complement</i>
Él me dice que	al _N secretario _{N+1} (que)	cuándo le _N ofrezco _{N+1} ...
Ella me pregunta que	esa _N casita _{N+1} de mi abuelo (que)	cuándo la _N pinto _{N+1} ...

Section two, the left-dislocated topic, begins immediately after the primary complementizer (C1). The critical region (N) is composed of the demonstrative adjective *ese/esa* or *al*, the contraction of preposition (*a*) and determiner (*el*). While these words are of similar length (2 or 3 characters), they do not share grammatical equivalency. Thus, we ran an independent samples t-test to determine whether reading times differed across items. We found no statistical significance between the two categories items, $t(22) = 0.759, p = 0.455$. Given that integration effects can “spill over” from the critical region, we analyze reading times here as well. The spillover region (N+1) is a noun of + or - animacy, controlled for lexical frequency, number of characters and syllable count (see Table 4.4). In the long intervener condition only, there is an additional prepositional phrase immediately following the aforementioned noun phrase region of interest. The noun within

the prepositional phrase is also controlled for lexical frequency, number of characters and syllable count (see Table 4.4).

The third and final sentence section, the *wh*-complement, is the primary section of interest for this study. As mentioned, we are first and foremost interested in determining whether an expectation-based or memory-based account of recomplementation best account for the measured effects of complement integration. The critical region (N) is a *wh*-word across all conditions, either *dónde* ‘where’ or *cuándo* ‘when’. The spillover regions (N+1 and N+2) are composed of the third person singular direct or indirect object pronouns *lo/la/le* and a transitive verb conjugated in the 1st person singular, respectively. The transitive verb was controlled for lexical frequency, number of characters and syllable count (see Table 4.4). Lastly, the sentence concludes with a three-word prepositional or noun phrase, which serves the sole purpose of accounting for wrap-up effects. All sentences were carefully controlled for length (13-16 words), where variation is due to the variables of intervener length and C2 lexicalization. A summary of the content words, which were controlled for lexical frequency, number of characters and syllable count is offered in Table 4.4.

Table 4.4. Content-word metrics

Content Words	<i>N</i>	<i>M</i>	<i>SD</i>
<i>Section 2: N+1 region, noun</i>			
Frequency per million	48	45.02	54.08
Number of characters	48	7.33	1.56
Number of syllables	48	3.44	0.50
<i>Section 2: long condition, noun2</i>			
Frequency per million	48	92.19	96.21
Number of characters	48	7.29	1.40
Number of syllables	48	3.17	0.38
<i>Section 3: N+2 region, verb</i>			
Frequency per million	48	60.40	48.78
Number of characters	48	6.48	1.38
Number of syllables	48	2.65	0.48

Content-word metrics were searched for in Davies’ *The Corpus del Español NOW* (Davies, 2017). This corpus is composed of news on the internet from 2012 to present day and contains more than 7.2 billion words. Lexical frequency is reported in a standard measure p frequency per million. It is defined as the total number of times the word appears in Davies’ corpus divided by the total count of the words in the corpus, then multiplied by one million. The nouns were either 3 or 4 syllables and only two of them were repeated more than once, while the verb was 2 or 3 syllables and 10 of them were repeated more than once.

As mentioned earlier, 48 master test items were created for the purpose of this task, 6 stimuli for each of the eight conditions. Following best practice in methodological design, each of the 48 master test items were manipulated by verb type, intervener length and C2 lexicalization so that each master item had a representative sample in each of the 8 conditions (Marull, 2017ab; Sagarra & Herchensohn, 2013). This resulted in 384 test items. Using Latin square design, these items were then distributed along eight lists such that one

condition alone from the master compilation was distributed to a single list. This process is summarized in Table 4.5 below and the master list of experimental stimuli is offered in Appendix H:

Table 4.5. Latin-square design

	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Item 8	...	Item 48
List 1	1a	2b	3c	4d	5e	6f	7g	8h	...	8h
List 2	1b	2c	3d	4e	5f	6g	7h	8a	...	8a
List 3	1c	2d	3e	4f	5g	6h	7a	8b	...	8b
List 4	1d	2e	3f	4g	5h	6a	7b	8c	...	8c
List 5	1e	2f	3g	4h	5a	6b	7c	8d	...	8d
List 6	1f	2g	3h	4a	5b	6c	7d	8e	...	8e
List 7	1g	2h	3a	4b	5c	6d	7e	8f	...	8f
List 8	1h	2a	3b	4c	5d	6e	7f	8g	...	8g

The forty-eight items in each list were then combined with forty-eight distractors of comparable length targeting unrelated linguistic phenomena (i.e., adverb and negative particle placement and agreement morphology). The stimuli were ordered using a pseudo-randomization design. First the items were randomized using Excel’s RANDBETWEEN function. The results were then corrected manually so that no test items appeared back-to-back.

Lastly, counterbalanced yes/no comprehension questions preceded each test item to ensure attention to task and to help distract from the primary goal of the experiment. The comprehension questions were meaning-based and either addressed gaps in knowledge (e.g., *Él sabe cuándo voy a escribir la recomendación* ‘He knows when I will write the recommendation’) or the specific items that were referenced (e.g., *Él habla de cuándo voy a servir la comida* ‘He speaks about when I will serve the food’). The self-paced reading experiment took on average 20 minutes to complete.

4.5 RESULTS

4.5.1 Pre-Analysis

Results from the self-paced reading (SPR) comprehension question were used to trim the data prior to analysis of reading times. As previously mentioned, it is common in SPR analysis on native speakers to remove all test items associated with incorrect comprehension question responses (Jegerski, 2014). Further, participants who score below a 70% on overall sentence comprehension are removed from the experiment altogether (Tucker et al., 2015). The underlying assumption here is that incorrect responses reflect a lack of attention to task. The outcome of this adjustment to the data is summarized in Table 4.6 below:

Table 4.6. Comprehension question results

Test group	N	M	SD	Range
Initial	24	0.843	0.127	0.537 to 0.979
Final	21	0.88	0.077	0.707 to 0.979

As demonstrated in Table 4.6, the mean number of correct responses of the initial test group (n=24) is 0.843 (84%). Three participants scored below a 70%, 0.658, 0.536 and 0.541 respectively, and were removed from the analysis. This results in an adjusted overall mean of 88%, or the removal of 12% of the remaining test items. Since linear mixed-effects modeling was adopted for primary analysis, no further data trimming of outlier reading times is required (Baayen, Davidson, & Bates, 2008; Baayen & Milin, 2010). As explained in Jegerski (2014), “given that mixed-effects models do not rely on aggregate means, the full range of values remains in the data on which the statistical tests are conducted and the presence of outliers is thus not such a concern” (20).

4.5.2 Matrix Verb Analysis

Our first research question considers whether the onset of the left-dislocated material is more difficult to integrate after the verb *preguntar* ‘to ask’ or *decir* ‘to say/tell’ (see Table 4.7).

Table 4.7. Region of interest for matrix verb analysis

Section 1	Section 2	
	<i>N</i>	<i>N+1</i>
Ella me dice que	esa	receta
Ella me pregunta que	esa	receta

Recall, we hypothesized that if verb biases are used in an expectation-based account to anticipate new information or predict structure, then processing complexity will be greater in critical region (N) and possibly spillover region (N+1) after the verb *preguntar* than *decir*. This is motivated by the fact that the ask/wonder verb *preguntar* obligatorily selects for a *wh*-complement, a condition which is not met at this incremental stage. While this is not our primary area of interest in this study, we offer a preliminary analysis with descriptive results summarized in Table 4.8.

Table 4.8. Descriptive statistics for matrix verb analysis

Verb	N (ms)		N+1 (ms)	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Decir	395.04	74.95	481.22	248.91
Preguntar	435.52	45.96	490.42	342.23
<i>Overall mean</i>	415.28	60.45	485.81	295.21

The descriptive statistics point to a trend in support of our hypothesis. Namely, integration of critical region (N) is more costly or is associated with a relatively higher reading time in the *preguntar* condition than the *decir* condition (435.52ms versus 395.04ms, respectively). This effect, however, does not appear to spillover into the N+1 region, where higher overall reading times associated with integration of longer and content-related words do not appear to differ by matrix verb type (490.42ms and 481.22ms, respectively).

In order to determine whether this trend is statistically significant, a linear mixed model was run with the LMER function in R (R Core Team, 2017). The model defined one fixed effect of verb (*decir* or *preguntar*) and one random intercept for subject. Degrees of freedom were estimated using Satterthwaite's method to calculate a *p*-value. As a result, we report *t*-values directly instead of *z*-scores (Tucker et al., 2015). The effect of verb type was not found to be significant in either the N ($\beta = 40.34, t(816.98) = 1.639, p = 0.101$) or N+1 ($\beta = -6.719, t(816.18) = -0.293, p = 0.77$) regions. Thus, we cannot claim with statistical evidence that the parser uses verb biases during the initial computation.

4.5.3 Complement Integration Analysis

Our primary area of interest in this study pertains to whether a memory-based or expectation-based model of recomplementation better accounts for the shape of the interaction between intervener length and C2 lexicalization. Casasanto and Sag (2008) provide the only experimental evidence to date of the shape of this interaction. In order to provide a second source of evidence, we investigated reading times associated with the integration of the *wh*-complement in recomplementation questions, see Table 4.9.

Table 4.9. Region of interest in *wh*-complement analysis

Section 1	Section 2		Test variable	Test variable	Section 3		
	<i>N</i>	<i>N+1</i>	<i>extra material</i>	<i>C2 lexicalization</i>	<i>N</i>	<i>N+1</i>	<i>N+2</i>
Ella me dice que	esa	receta	(prepositional phrase)	(que)	cuándo	la	cocino...
Ella me pregunta que	esa	receta	(prepositional phrase)	(que)	cuándo	la	cocino...

Specifically, we analyzed the critical region (N) and the potential of two spillover regions (N+1) and (N+2) in section 3. Importantly, the N+2 region is not the end of the sentence. In order to account for wrap-up effects, all test items concluded with a three-word noun or prepositional phrase. It is also important to note that extra intervening material and C2 lexicalization are operationalized as experimental variables of interest. Table 4.10 offers the descriptive statistics for our analysis. The matrix verb has been collapsed for ease of interpretation, as well as due to a lack of significant effect on reading times as found in the matrix verb analysis (see Section 4.5.2).

Table 4.10. Descriptive statistics for *wh*-complement integration analysis

Condition	Intervener Length	C2	N (ms)		N+1 (ms)		N+2 (ms)	
			<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
1	Short	Overt	498.77	302.36	418.06	168.98	468.9	252.58
2	Short	Null	496.27	277.50	416.39	177.74	501.42	347.09
3	Long	Overt	437.36	201.04	385.04	141.65	470.03	286.80
4	Long	Null	486.25	330.14	393.44	162.65	463.8	378.78
		<i>Overall</i>	479.66	273.76	403.23	158.57	476.04	312.60

Recall that the critical region (N) represents the *wh*-word, spillover region 1 (N+1) represents the object pronoun, and spillover region 2 (N+2) represents the finite verb. Thus, it is not surprising that the object pronoun would on average be read faster than the other two regions, which contain longer and content-related words. Of greater interest, we see a trend, where the long intervener conditions are read faster than the short intervener conditions. This trend is clear across all three regions of interest. This is surprising under a distance dependency theory of processing complexity. On the contrary, distance appears to have a facilitation effect. Furthermore, in the critical region (N) there also appears to be an interaction between length and C2, where overt C2 serves a facilitative function in the long condition but not the short.

In order to investigate these observed trends, we ran three linear mixed models with the LMER function in R (R Core Team, 2017). The models were run for the N, N+1 and N+2 regions and all defined three fixed effects of matrix verb (*decir* or *preguntar*), intervener length (short or long) and C2 lexicalization (overt or null), an interaction of intervener length*C2 lexicalization and one random intercept for subject. Even though matrix verb does not directly inform our analysis here, we chose to keep it as a fixed effect in order to not overlook the possibility of downstream effects that were not caught in our first analysis. As before, degrees of freedom were estimated using Satterthwaite's method to calculate a *p*-value. As a result, we report *t*-values directly instead of *z*-scores (Tucker et al., 2015). Statistical results are summarized in Table 4.11.

Table 4.11. Summary of linear mixed effects analysis

Region	Matrix verb	Intervener length	C2 lexicalization	Length*C2
Critical Region (N)	$\beta = 18.19$, $t(813.23) = 1.089$ $p = 0.276$	$\beta = -63.16$, $t(813.74) = -2.689$ $p = 0.007$	$\beta = -0.506$, $t(814.74) = -0.022$ $p = 0.982$	$\beta = 61.38$, $t(813.77) = 1.833$ $p = 0.067$
Spillover 1 (N+1)	$\beta = 6.296$, $t(813.26) = 0.692$ $p = 0.489$	$\beta = -34.65$, $t(813.66) = -2.708$ $p = 0.006$	$\beta = -1.398$, $t(814.48) = -0.109$ $p = 0.913$	$\beta = 16.23$, $t(813.69) = 0.889$ $p = 0.374$
Spillover 2 (N+2)	$\beta = -33.95$, $t(813.17) = -1.372$ $p = 0.170$	$\beta = -16.11$, $t(813.47) = -0.911$ $p = 0.362$	$\beta = -6.402$, $t(813.79) = -0.253$ $p = 0.800$	$\beta = 53.91$, $t(813.36) = 1.526$ $p = 0.127$

Results from the battery of statistical tests largely confirm noticeable trends in the descriptive analysis. Specifically, intervener length in both the N and N+1 regions was found to be significant. This suggests that a facilitation effect of length is not only centralized in the *wh*-word (N) but it also spills over to the object pronoun (N+1). The hypothesized interaction of Length*C2 in the critical region was not significant ($p = 0.067$). Since intervener length was significant in both N and N+1 regions and no significant effects or interactions were found in the N+2 region, we collapsed the reading times of regions N and N+1 and ran one last LMER with the same specifications as before. Intervener length was highly significant ($\beta = -97.88$, $t(813.46) = -3.446$, $p = 0.0005$) and the interaction of length*C2 approached significance ($\beta = 77.95$, $t(813.48) = 1.925$, $p = 0.054$). C2 ($\beta = -1.821$, $t(814.07) = -0.064$, $p = 0.948$) and verb ($\beta = 24.44$, $t(813.16) = 1.210$, $p = 0.226$) were not significant.

The fact that the interaction of length*C2 has been previously documented as significant (Casasanto & Sag, 2008), as well as the near significance in our findings motivated a post hoc analysis. We ran a pairwise comparison with Bonferroni adjustment with the emmeans function in R. Results showed that C2 lexicalization significantly

facilitated complement integration in the long intervener condition ($\beta = -76.14$, $t(814) = -2.626$, $p = 0.043$) but not in the short condition ($\beta = 1.82$, $t(814) = 0.064$, $p = 0.999$). This interaction is visually depicted in Figure 4.6, alongside Casasanto & Sag's (2008) results reproduced in Figure 4.7.

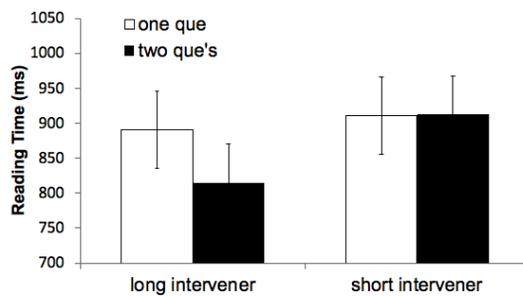


Figure 4.6. Present study self-paced reading results

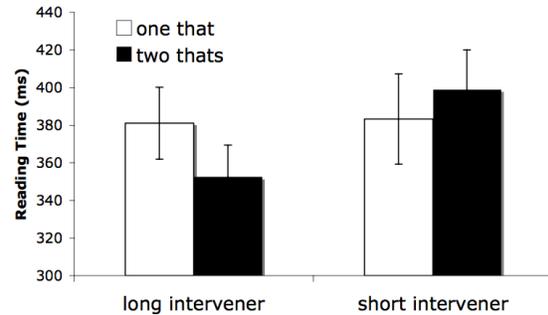


Figure 4.7. Casasanto and Sag self-paced reading results

Note the different y-axis scales in Figures 4.6 and 4.7. The overall reading times are much higher in the former owing to the collapsed N and N+1 regions. Scale aside, the patterns across both graphs are very similar. As mentioned in Section 4.3.1, Casasanto and Sag (2008) only report the statistical results of the interaction and their post hoc analysis, where the long condition was significant and the short condition was not. This makes it difficult for us to compare experimental results much further. However, our statistical analysis with a significant effect for intervener length, as well as the full post hoc pairwise comparison, suggest that there is more to the story. The full pairwise comparison is shared in Table 4.12:

Table 4.12. Pairwise comparison post hoc analysis

Contrast	Results	$p < 0.05$
overt/short - null/short	$\beta = 1.82, t(814) = 0.064, p = 0.999$	No
overt/long - null/long	$\beta = -76.14, t(814) = -2.626, p = 0.043$	Yes
overt/short - overt/long	$\beta = 97.88, t(813) = 3.446, p = 0.003$	Yes
overt/short - null/long	$\beta = 21.74, t(814) = 0.759, p = 0.872$	No
null/short - overt/long	$\beta = 96.06, t(814) = 3.341, p = 0.004$	Yes
null/short - null/long	$\beta = 19.92, t(813) = 0.691, p = 0.900$	No

We will take each result in turn. The first two rows have already been reported on and refer to the lack of significant effect for overt C2 in the short condition and the significant effect in the long condition. In Casasanto and Sag (2008), this pair of results serves as the basis for their *advantage of the ungrammatical* hypothesis. To be a bit more specific, the authors argue that there is a trend for the null/short condition to be easier to process than the overt/short one. Importantly, we find no such trend in our results. In row three, we see a significant effect, where the complement is integrated faster in the overt C2 condition when the intervener is long as compared to short. This result is compatible with the DLT+ hypothesis. After all, if overt C2 is associated with an unspecified grammatical violation, then in contexts where C2 lexicalization provides no processing benefit (i.e., the short condition) an anomalous C2 might increase processing complexity. In row four, we see no significance between the overt/short and null/long conditions. Adopting Casasanto and Sag's terminology, there is no significant difference in the processing complexity of (1) a grammatical violation that minimally reduces dependency distance and (2) the absence of a violation in a long-distance dependency. This comparison is not particularly insightful for the purpose of the present chapter.

Given the post hoc pairwise comparison up to this point (i.e., rows 1-4 in table 12), we might be compelled to agree with the processing/grammaticality tradeoff theory or DLT+. However, we did find a highly significant main effect for length, which is difficult to reconcile. The remainder of the pairwise comparison provides deeper insights. Row 5 demonstrates that processing complexity is higher in the null/short condition than the overt/long one, while row 6 shows no significant difference in the complexity of the null/short and the null/long conditions. These two results are surprising given a processing/grammaticality tradeoff theory. First, according to the theory, null/short items display no grammatical violation and a short distance dependency. If overt C2 violates the grammar, then it's difficult to reconcile why the overt/long condition is significantly easier to process than the null/short one. Secondly, DLT and DLT+ models easily predict that the null/short condition would be read significantly faster than the null/long one. However, no significance is found. In Table 13, we compare key similarities and differences in the reported findings of Casasanto and Sag (2008) and the present study.

Table 4.13. A comparison of results, Casasanto and Sag and the present study

Selected Statistical Report	C&S (2008)	Present Study
Significant interaction of C2*length	yes	$p = 0.054$
Overt/long condition significantly less complex than null/long	yes	yes
(Insignificant) trend that null/short is easier to process than overt/short	yes	no
Overt/long condition significantly less complex than null/short	not reported	yes
Null/short condition is significantly less complex than null/long	not reported	no
Main effect for length	not reported	yes

In summary, despite similar overall patterns in Figures 4.6 and 4.7, the presented data analysis brings the *advantage of the ungrammatical hypothesis*, or the DLT+ account, into question. Before we consider whether the data is more compatible with an experienced-based account, we review one last experimental variable. Specifically, we supplement our primary study with an exploratory investigation of whether individual differences in working memory span predict performance.

4.5.4 Working Memory Analysis

Recall, for research question three we hypothesized that a DLT or DLT+ account of recomplementation would predict that comprehenders with relatively lower working memory span benefit most from C2 lexicalization in long distance dependencies. Dependency distance equals complexity and those with less working memory capacity have fewer available resources. On the other hand, an expectation-based account predicts the comprehender with higher working memory capacity benefits most from C2 lexicalization. Specifically, this group has more resources to allocate to predicting the C2 complexity correlate, which becomes increasingly more likely as a function of distance. When expectations are met, reading times are faster. When expectations are not met, reading times are slower.

In order to test these predictions, we ran a series of Pearson correlations to determine whether working memory span predicted reading times in several long intervener scenarios. With two exceptions, reading times were measured as a sum of N “*wh*-word” and N+1 “object pronoun” regions. Exception one pertains to a correlation we ran on reading times for the entire sentence. Exception two pertains to a correlation we ran for the secondary *que* itself (N-1). The battery of correlations is summarized in Table 4.14.

Table 4.14. Summary of correlations with working memory

Condition	Predictor	<i>r</i>	<i>p</i> value	Interpretation
entire sentence	working memory	-0.348	0.061	weak and insignificant negative correlation
null/long	working memory	-0.335	0.068	weak and insignificant negative correlation
overt/long	working memory	-0.475	0.014	moderate and significant negative correlation
null - overt long	working memory	0.186	0.209	very weak and insignificant positive correlation
short - long	working memory	0.105	0.324	very weak and insignificant positive correlation
<i>que</i> region short - long	working memory	0.482	0.013	moderate and significant positive correlation

In order to establish a baseline, we first investigated whether working memory predicted the reading time of the entire sentence. We found a weak and insignificant negative correlation ($r = -0.348$, $p = 0.061$). We interpret this as a lack of statistical evidence that comprehenders with higher working memory span have faster (i.e., lower) reading times in general. This result establishes that any significant correlations cannot be solely explained by the general phenomenon that those with higher capacity read faster. Of the remaining correlations, two reached significance. A moderately strong and significant negative correlation was found in the overt/long condition ($r = -0.475$, $p = 0.014$) and a moderately strong and significant positive correlation was found in the *que* region for the short minus long conditions ($r = 0.482$, $p = 0.013$). We depict the correlations and confirm assumptions of linearity and directionality in Figures 4.8 and 4.9.

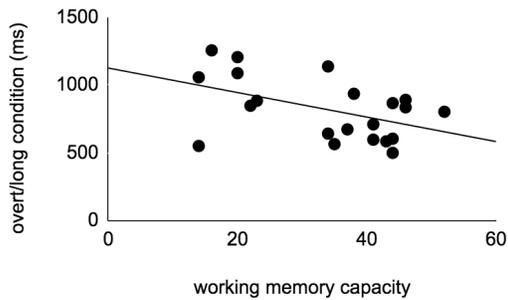


Figure 4.8. Negative correlation between working memory and overt/long condition

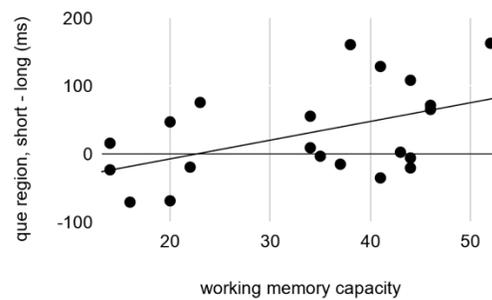


Figure 4.9. Positive correlation between working memory and *que* region, short-long conditions

The negative correlation between working memory capacity and reading times in the overt/long condition suggests that participants with a higher working memory capacity displayed faster (lower) reading times. We hypothesized that this outcome would support the expectation-based account and would be surprising given a memory-based account. The interpretation of the positive correlation between working memory capacity and reading times in the *que* region, short minus long conditions, is a little less straightforward. Specifically, as working memory capacity increases, so do the reading times of the secondary *que* in the short condition relative to the long one. This result provides supplemental evidence in support of an expectation-based account, where information that is expected— overt C2 in the long condition— is associated with faster (lower) reading times. DLT and DLT+ accounts don't make strong predictions about the reading time of the complementizer, which simply serves as a bridge for the dependency relationship between the matrix verb and the complement (Casasanto & Sag, 2008). Specifically, at the time of complementizer integration, the dependency relationship has not yet been resolved. DLT+ might offer that the secondary *que* in the short condition would be read more slowly across all comprehenders owing to a grammatical violation with little benefit to processing.

However, we can derive no clear predictions from memory-based accounts about the secondary *que* in the long condition.

4.6 DISCUSSION

The goal of the present chapter was to shed light on whether expectation-based or memory-based proposals of syntactic parsing best account for the phenomenon of recomplementation. In order to accomplish this goal, we first reviewed the predictions made by two variations of Dependency Locality Theory *vis-à-vis* the shape of their respective interactions. We labeled these accounts as the DLT and the DLT+ models, where the former strictly follows the tenets of Dependency Locality Theory and the latter adopts a grammar/processing tradeoff theory. The DLT+ account, as introduced by Casasanto and Sag (2008), proposes that C2 lexicalization is a violation of the grammar, which accounts for an insignificant trend in Casasanto and Sag's data in the direction of the overt/short condition being more difficult to process than the overt/null condition. A stricter account of Dependency Locality Theory makes no such prediction and simply adheres to the principle that complexity is equal to the distance between two items engaged in a syntactic dependency. We then compared Casasanto and Sag's experimental results to the two models and concluded that neither DLT nor DLT+ fit the data particularly well. Moreover, the stipulations of DLT+ are not compatible with the syntactic-theoretical literature on the grammaticality of the secondary complementizer and its multifunctionality.

This motivated us to propose an expectation-based account that better fit Casasanto and Sag's data and was compatible with the formal literature. Grounded in the literature that converges on the positive correlation between C2 lexicalization and intervener length, we proposed that secondary *que* is a complexity correlate that *can* be probabilistically

predicted along the intervener length constraint. Further, we hypothesized that given a highly predictive parallel parser, the complexity correlate *would* be predicted as a function of distance. At this point in the chapter, we largely remained agnostic as to whether expectation-based or memory-based models better explain the phenomenon of recomplementation. We simply argued that the former better explained Casasanto and Sag's data.

In order to make a more informed decision, we provided new experimental evidence. We adopted Casasanto and Sag's self-paced reading paradigm to offer a second experimental study that investigates complement integration in recomplementation constructions. In addition, we chose to investigate question sentence types in order to offer preliminary evidence of the effect of matrix verb biases in the incremental processing of recomplementation. Lastly, we offered preliminary evidence on the relationship between individual differences in working memory capacity and performance. Three specific research questions drove our experimental analysis: (RQ1) Are matrix verb biases used to anticipate new information?; (RQ2) Do memory-based or expectation-based models better account for the shape of the interaction between intervener length and C2 lexicalization?; and (RQ3) What is the relationship between individual differences in working memory capacity and online performance?

For (RQ1), we hypothesized if verb biases are used in an expectation-based account to anticipate new information, then processing complexity will be greater after the verb *preguntar* than *decir*. This hypothesis was grounded in the selectional properties of each verb. For (RQ2), our primary area of investigation, we offered a series of hypotheses that naturally fall from the memory-based and expectation-based accounts (as reviewed in Section 4.2.1). These hypotheses included but were not limited to divergent predictions between DLT and DLT+ accounts with respect to overt/short conditions, as well as

divergent predictions between DLT and DLT+ on the one hand and our expectation-based account on the other, with respect to the processing complexity in short versus long conditions more generally. (RQ3) supplemented our primary area of investigation with an analysis at the level of individual differences. We hypothesized that in long intervener conditions, DLT and DLT+ models predict that individuals with relatively lower working memory benefit most from C2 lexicalization, while the expectation-based account predicts that individuals with higher working memory capacity benefit the most.

Experimental results from our self-paced reading task are much better aligned to the expectation-based account of recomplementation and the shape of the interaction it predicts (see Figure 4.4), as opposed to either memory-based accounts. Our results are summarized in Table 4.15:

Table 4.15. Evidence for memory-based versus expectation-based accounts

Experimental evidence	Memory-based account		Expectation-based account
	<i>DLT</i>	<i>DLT+</i>	
matrix verb biases used to anticipate	n/a	n/a	inconclusive
main effect: intervener length	×	×	✓
significant interaction: C2*length	✓	✓	✓
working memory capacity	×	×	✓
<i>shape of the interaction</i>			
overt/long condition < null/long condition	✓	✓	✓
overt/short condition = null/short condition	×	×	✓
null/short condition = null/long condition	×	×	✓
overt/short condition > overt/long condition	×	✓	✓

One advantage of adopting a forward-looking account of recomplementation is we can start to align the incremental processing of recomplementation with the extensive literature on the role of prediction in syntactic parsing (for a review see Kuperberg & Jaeger, 2016). For example, verb biases have been shown to influence prediction of new information and structural hypotheses (e.g., Altmann & Kamide, 1999, 2007; Kamide et al., 2003). In the present study, we found a trend in favor of our hypothesis that a [-wh] phrase like *esa receta* ‘that recipe’, would be more difficult to integrate when introduced by an ask/wonder verb [+wh] as opposed to a non-ask/wonder verb [-wh]. However, investigation is only preliminary and the trend was not statistically significant.

In our primary analysis pertaining to *wh*-complement integration, we did find a main effect for intervener length and a significant interaction of C2 lexicalization and intervener length. The main effect of intervener length suggests that distance reduces processing complexity. This is difficult to reconcile given a DLT model. Further, it’s unclear the extent to which the DLT+ model can account for this effect by turning to grammar/processing tradeoff theory. DLT+ does clearly account for overt/short condition being more difficult to process than the overt/long condition, but it does not account for the equivalent difficulty of the null/short and null/long conditions. The expectation-based account, on the other hand, adopts a probability distribution of predicted structures framework. Recomplementation as a left-dislocated phenomenon, is an infrequent, non-canonical form. Thus, we hypothesize that distance (or time) will facilitate the appropriate distribution reweighting, from a more frequent hypothesis (single complementizer) to a less frequent one (double complementizer). Given that the shift in probability distribution is associated with processing cost, we believe that an expectation-based model can more easily account for the main effect of length.

The interaction between C2 lexicalization and intervener length suggests that the overt/long condition is the least complex condition. This is consistent across all three models. In review, the memory-based models argue that lexicalized C2 reduces processing complexity by erasing the dependency distance. The expectation-based model argues that C2 is a complexity correlate that is probabilistically predicted as a function of distance. When expectations are met, processing complexity goes down. When expectations are not met, processing complexity increases. Where the predictions of the models differ is with respect to the other three conditions (see Figures 4.2-4.4). Memory-based models fail to accurately predict the equivalent processing complexity of the overt/short, null/short and null/long conditions. We give the slight nod to DLT+ over DLT for accurately predicting two out of the four total conditions. The expectation-based model, on the other hand, accurately accounts for all four conditions. Importantly, we find the results from our study to be wholly consistent with Casasanto and Sag's data, where even their results in our view are best explained by an expectation-based account (for discussion see Section 4.2.1).

Lastly, we supported our primary analysis with evidence from individual differences in working memory capacity. Working memory was found to negatively correlate with reading times in the overt/long condition and positively correlate with reading times in the *que* region for the short minus long conditions. We interpret faster reading times in the overt/long condition for comprehenders with more available resources as evidence of a facilitation effect associated with expectations met. Further, we interpret the trend for comprehenders with more resources to read secondary *que* faster in the long as compared to the short conditions as supplemental support for an expectation-based account. C2 lexicalization is probabilistically predicted as a function of distance. In the same token, it is not expected when intervener length is short. Memory-based accounts on the other hand, offer us very little in the way of reading time predictions for the secondary

‘She asks me when I will cook the chef’s recipe’. The sentence has been divided into four sections for ease of description. Section 1, or input 1, pertains to the matrix verb. If, as hypothesized, verb biases inform syntactic parsing, then we expect the canonical indirect question hypothesis to be ordered first in the probability distribution. In our model, we do not commit to the number of representations that can be activated in parallel. For ease of exposition, we display two, with the uppermost corresponding to the structure with the highest degree of probability. Input 2 pertains to the onset of the left-dislocated material. Because the [+wh] feature of the matrix verb doesn’t match the [-wh] feature of the ensuing phrase, the parser shifts the ordering of the probability distribution, now favoring a hypothesis that can account for present evidence. Input 3, which corresponds to a longer intervener length or extra material in the left-dislocation, reinforces the ordering of hypotheses. This results in a strengthening of the commitment to the most probable representation (e.g., 0.70/0.30 probability becomes 0.80/0.20). Crucially, the parser also predicts the complexity correlate, or C2 lexicalization, with a strength that increases as a function of distance. Lastly, for input 4, which corresponds to *wh*-complement integration, the parser strengthens the commitment even further, assuming the expectation of overt C2 is met. If not met, probability distribution is reweighted and costs of integration increase.

In conclusion, we believe the incremental steps in our expectation-based model offer a coherent, data-driven account of the online comprehension of recomplementation. Since we have only accounted for a subset of recomplementation sentence types in our experiment, the model should be further tested on other variations, including but not limited to statements, different moods, as well as intervener clause types and lengths. Future research should also continue to investigate the effect of verb biases on hypothesis formation, as well as the effect of individual differences on parsing strategy. We only begin to explore these two variables in this chapter. Lastly, given that recomplementation is

largely considered an oral phenomenon in present-day Spanish, it will be important to investigate the effect of intonation on the integration of new information.

Chapter 5: Conclusion

The narrow goal of this dissertation was to address limitations in the extant literature on recomplementation. These shortcomings primarily pertain to linguistic population biases (e.g., the Peninsular bias), methods and analysis, and the gap between syntactic-theoretical accounts and experimental findings. The broader aim was to provide a blueprint or checklist for how to address these failings towards advancing the field of syntax more generally. When proposing theoretical accounts or models, the researcher can and should consider the following questions:

- (i) Have speakers with diverse profiles (e.g., heritage speakers) informed general theory?
- (ii) Have understudied dialects been considered (e.g., the Caribbean lect)?
- (iii) Are experimental findings, psycholinguistic models and syntactic-theoretical accounts aligned?

The dissertation goals and checklist items were addressed by offering three experimental recomplementation studies. Study 1 specifically investigated heritage speaker grammar via aural acceptability judgment and forced-choice preference tasks. Study 2 explored microvariations in Colombian and Cuban Spanish through the offline production tasks of elicited imitation and sentence completion. Finally, study 3 analyzed the incremental processing of recomplementation sentences by way of self-paced reading paradigm. It further explored individual differences in working memory as a predictor of performance.

5.1 SUMMARY

The three primary research questions that drive this dissertation, along with the corresponding sub-questions from each study, are reproduced below. In addition, each set

of questions is accompanied by a review of the results and a discussion of the broader contributions.

- (RQ1) Do advanced heritage speaker grammars diverge from a relevant baseline? If so, how does bilingual data inform syntactic-theoretical accounts of recomplementation?
- (RQ1.1) Do advanced heritage speakers accept the null C2 construction at a higher rate than the overt C2 option? Does language use or proficiency predict this outcome?
- (RQ1.2) Do advanced heritage speakers prefer the null C2 construction at a higher rate than the overt C2 option? Does language use or proficiency predict this outcome?
- (RQ1.3) In terms of (RQ1.1) and (RQ1.2), do advanced heritage speakers diverge from the monolingual baseline group?

With respect to (RQ1), it was hypothesized that the heritage speaker group would diverge from the baseline. Specifically, they would accept and prefer the overt C2 at a higher rate than the control group. This hypothesis was grounded in previous research on the vulnerability of CP-related phenomena in bilingual populations and interpreted within the framework of Polinsky and Scontras' (2020) model of divergent attainment. Specifically, null elements and distance dependencies are common sources of divergence between heritage grammars and a relevant baseline. The results largely confirm the hypothesis. While the baseline group accepts the null variety at a significantly higher rate, heritage speakers display no significant effect for C2 lexicalization. This result is supported by an individual analysis that shows that bilingual participants who rate one variety as more acceptable than the other do so by only a small margin. In the supplemental preference task, expectations of divergent behavior are also confirmed. The baseline group prefers the null variety across both question and statement sentences at a significantly higher rate than

the heritage group. Curiously, a significant interaction of group and sentence type shows that the heritage speakers' proportion of C2 lexicalization preference is significantly greater in questions than in statements. In summary, the divergent behaviors of the two groups can be summarized as a marginal to no effect of C2 in the heritage speaker group and a significant effect of C2, favoring the null variety, in the baseline group.

In broader terms, we discuss how speakers with diverse profiles can inform general theory and perspectives on the nature of grammatical representations and linguistic complexity. The avoidance of silent phenomena refers to the greater use of overt varieties in some linguistic populations when compared to others. With respect to pronominal expression, some of these populations are in a situation of contact (e.g., de Prada Pérez, 2009), while others seem to avoid silent phenomena as characteristic of a regional variety (e.g., Camacho, 2013; Lipski, 1977; Martinez-Sanz, 2011; Orozco, 2015; Ortiz-López, 2009; Otheguy & Zentella, 2012; Toribio, 2000). For example, traditional assumptions of pro-drop do not hold in the Caribbean lect. Importantly, avoidance of silence in either bilingual or monolingual varieties does not imply the absence of null forms. Rather, the proportion of null cases is lower when compared to a relevant baseline. Alternatively, the Spanish variety to which they are exposed in early childhood may optionally select for null and overt varieties in the given conditions. As noted by several scholars, the adoption of overt forms can be traced to earlier generations (e.g., Montrul, 2016; Otheguy & Zentella, 2012; Otheguy et al., 2007; Sorace, 2004). Thus, as is characteristic of all linguistic populations, speakers acquire the language they are exposed to.

What's more, according to Casasanto and Sag's (2008) memory-based model of recomplementation, the high complementizer (C1) predicts a complement (e.g., embedded subject or embedded verb in the case of pro-drop languages like Spanish). This dependency remains unresolved until complement integration, at which point old information must be

retrieved. C2 lexicalization (as a reiteration of C1, see DoubledForceP account) reduces the length of the dependency or the retrieval distance to zero, lowering the strain on working memory prior to complement integration. Given that the strain on available resources should be even greater in bilinguals and heritage speakers specifically, who are holding multiple languages in parallel while working in the less dominant one, it is not surprising that they favor forms that reduce processing complexity (e.g., Keating et al., 2016; Montrul, 2016; Polinsky & Scontras, 2019; Sánchez, 2019). Further, we would predict that individuals with fewer available resources (i.e., lower working memory capacity), in general, would also prefer the overt C2 at a higher rate when compared to individuals with a higher working memory span. Thus, we propose a multiple representations account of recomplementation, where processing complexity triggers divergent attainment. Namely, heritage speakers develop a DoubledForceP abstract representation of the phenomenon under consideration.

In summary, speakers with diverse profiles, like speakers from all language varieties, make an important contribution to theory. Heritage speaker populations, for example, contribute to our general understanding of the role of experience and individual cognitive traits on language development and grammatical representation. In this dissertation, we argue that just as an understudied monolingual Caribbean lect informed our understanding of Spanish pro-drop (e.g., Camacho, 2013; Lipski, 1977; Martinez-Sanz, 2011; Orozco, 2015; Ortiz-López, 2009; Otheguy & Zentella, 2012; Toribio, 2000), US heritage Spanish informs our theory of recomplementation. Given several well documented domains of divergent attainment and relatively high within group variation, heritage languages stand to contribute to the theoretical development of many linguistic phenomena.

- (RQ2) Is recomplementation a locus of dialectal variation? If so, does this help to explain the divide in theoretical accounts and experimental findings?
- (RQ2.1) Is overt C2 licensed by Colombian Spanish grammar unlike in Cuban?
- (RQ2.2) Is recomplementation a locus of dialectal variation, microvariations considered?
- (RQ2.3) Does overt C2 facilitate complement integration (e.g., does intervener length predict secondary *que* expression)?

With respect to (RQ2), it was hypothesized that recomplementation was indeed a locus of morphosyntactic variation in Cuban versus Colombian Spanish. This hypothesis is grounded in Frank and Toribio's (2017) findings that recomplementation is not licensed by Cuban Spanish grammar juxtaposed with extant experimental and syntactic-theoretical literature arguing for its grammaticality in Peninsular Spanish (e.g., Villa-García, 2019) and 'Mainland' Spanish (e.g., Frank, 2016). The results did not confirm the hypothesis. Cuban and Colombian Spanish participants repeat null C2 condition test items with greater accuracy (94% and 99%) than the overt C2 condition test items (26% and 23%). Given the assumptions of an elicited repetition task, we preliminarily conclude that overt C2 is neither licensed by Spanish grammar nor is it a locus of dialectal variation. Lastly, we find no evidence that secondary *que* is adopted as a production strategy to facilitate complement integration. This finding is robust across dialects and all microvariations, including those related to intervener type, intervener length and sentence type.

Importantly, given a robust body of literature arguing for the grammaticality of recomplementation, the experimental results are met with criticism. Namely, the potential of task effect cannot be ruled out. For example, in the acceptability judgment task administered in study 1, a Colombian group rated overt C2 within the range of marginal acceptability (2.82 for questions and 2.99 for statements on a 1-totally acceptable to 7-

totally unacceptable scale). This more nuanced take on grammaticality might suggest that an elicited imitation task, with a binary dependent measure interpreted as licensed or unlicensed by the grammar, is not sensitive enough to investigate recomplementation. Further, potential conflation of production and comprehension effects are introduced in both tasks. Specifically, secondary *que* may not be produced in elicited imitation or sentence completion tasks because it only serves a facilitatory function in comprehension. Lastly, to more directly measure the benefit of C2 lexicalization on complement integration, an online methodology should be adopted. These concerns together bring into question the validity of the two offline tasks, making a strong conclusion with regard to recomplementation as a locus of dialectal variation difficult.

The broader research question considered the divide in theoretical accounts and experimental findings. We offer three potential explanations. First, it is possible that recomplementation is licensed by Spanish grammar and the tasks selected in the present study are inappropriate. The second explanation is that recomplementation is not licensed by Spanish grammar. Questions of task effect aside, the present study investigated Mainland versus Caribbean varieties of Spanish and found evidence that brings the grammatical status into question. This claim is further strengthened by the microvariational analysis across intervener length, intervener type and sentence type experimental items. The third explanation pertains to a well-documented researcher intuition or selection bias (Edelman & Christiansen 2003; Ferreira, 2005; Gibson & Fedorenko 2010ab; Ortega-Santos, 2020ab). With respect to recomplementation, the syntactic-theoretical literature has promoted a Peninsular bias. Given that the present study has not investigated Peninsular varieties, it is still too early to rule out recomplementation as a locus of dialectal variation. Importantly, we have added to the narrow list of experimental testing methods that have explored recomplementation. Further we advance the message that quantitative

experiments and non-quantitative syntactic-theoretical analyses must continue to inform one another, while paying close attention to such variables as dialect, microvariation, and testing method.

(RQ3) How can we reconcile psycholinguistic models of recomplementation with syntactic-theoretical accounts? Is it a worthy pursuit?

(RQ3.1) Are matrix verb biases used to anticipate new information?

(RQ3.2) Do memory-based or expectation-based models better account for the shape of the interaction between intervener length and C2 lexicalization?

(RQ3.3) What is the relationship between individual differences in working memory capacity and online performance?

In study 2, the gap between syntactic-theoretical accounts and quantitative findings remained unresolved. Further, questions around task selection loomed large. Study 3 accounts for these limitations. With respect to (RQ3), we hypothesized that an expectation-based model that is compatible with syntactic-theoretical accounts would better predict the shape of the interaction between intervener length and C2 lexicalization when compared to a memory-based model that is not compatible with theoretical accounts (Casasanto & Sag, 2008). The results confirm this hypothesis. The expectation-based model accounted for all four points of the interaction (i.e., overt/short, null/short, overt/long, null/long), whereas the memory-based model only accounted for two of the conditions. Further, in the overt/long condition exclusively, comprehenders with higher working memory capacity displayed faster reading times. We interpret this result as evidence in favor of the hypothesis. That is to say, comprehenders with more available resources are able to predict C2 lexicalization as a function of distance, which leads to a facilitation effect when expectations are met. Lastly, an exploratory analysis of the parser's ability to use verb

biases to anticipate new information also proves promising (e.g., Kamide et al., 2003), though future research is required to explore this trend in the data.

The broader research question pertained to whether building psycholinguistic models that are informed by syntactic theory is a worthy pursuit. With this study, we have provided an example where the answer is *yes*. By leading with the theory, we initially rejected Casasanto and Sag's (2008) proposal. Their model proposes a singular function of C2 lexicalization and the ungrammaticality of recomplementation. The proposal runs counter to the extensive body of literature that argues otherwise (e.g., Radford, 2013, 2018; Villa-García, 2015, 2019). This motivated us to propose an alternative model that was theory-driven. This exercise resulted in an expectation-based model that more accurately accounted for the data. As a secondary contribution, we showed that including an analysis at the level of individual differences offered a more comprehensive picture of the comprehenders' behaviors and provided even more evidence in support of the model.

In conclusion, in this dissertation we addressed shortcomings in the recomplementation literature that pertained to linguistic population biases and the divide between experimental findings and theoretical accounts. In so doing, we offered a roadmap for how to account for limitations in the field more generally. We argued that US heritage Spanish informs theoretical debates and should be treated as a language variety like Peninsular or Caribbean Spanish. We further advanced the message that quantitative experiments and non-quantitative syntactic-theoretical analyses alike should be informed by understudied dialects, should investigate linguistic microvariations, and should implement various methods (e.g., online vs. offline) and modes (e.g., oral vs. written) of data collection, while also replicating existing experiments. Lastly, we demonstrated that reconciling psycholinguistic models with syntactic theory is a worthy pursuit.

5.2 LIMITATIONS

As has been alluded to in the previous section, each of the three studies is not without limitations. Areas for improvement are experimental in nature and fall into three general categories: statistical power, experimental task and design, and participant selection. By statistical power, we refer to the possible inflation of type II error. We adopt the standard significance level of 0.05 to set a conservative criterion and avoid type I error. However, given relatively low participant numbers, this increases the chance of false negatives, or inconclusive nonsignificant effects. Experimental task and design refers to the type of task (e.g., online versus offline; production versus comprehension) and mode (e.g., oral versus written) that was selected, as well as the adopted fixed effects. Lastly, participant selection refers to the comparison of bilinguals to monolinguals.

In study 1, group sizes are at a bare minimum for the test and control groups, $n=15$ and $n=12$, respectively. We acknowledge that experimental power to detect differences was low and it is possible that both significant and nonsignificant effects are inconclusive. It is further acknowledged that test items do not account for a representative sample of the recomplementation phenomenon. Test items only vary by sentence type (i.e., question versus statement), where variations of intervener length, intervener type and mood are held constant. Other methodological issues pertain to the use of offline methods to make claims of processing benefit and the adoption of different modes in each task. Specifically, the acceptability judgment task was a listening experiment, while the preference task adopted the written mode. One last limitation of study 1 pertains to the monolingual control group. An ideal heritage speaker comparison group would not only be bilingual, but would also originate in the same community (e.g., see relevant baseline, Polinsky & Scontras, 2020). As a result, it is difficult to make strong claims with regard to the input as a trigger hypothesis.

Shortcomings of between-group comparison and representative sampling of recomplementation in study 1 are accounted for in study 2. However, limitations in this experiment must also be acknowledged. Again, we must concede that the number of participants is low in the Colombian (n=16) and Cuban (n=25) experimental groups. Further, the microvariational analysis that investigates different intervener types, intervener lengths and sentence types only adopts a small number of experimental items per condition. Given a low number of participants and experimental items, we must acknowledge the possible inflation of type I and type II error. Lastly, we consider the real possibility of task effect here. Specifically, the binary dependent measure of the elicited imitation task (i.e., accurate versus inaccurate repetition) may not be sensitive enough to investigate recomplementation. We also note a potential conflation of production and comprehension effects. Secondary *que* could in theory facilitate comprehension but not be produced in elicited repetition because it does not facilitate speaker production. Further, to more directly measure facilitation effects, an online methodology should be adopted. These concerns together bring into question the validity of the elicited imitation and the sentence completion offline tasks and make a strong conclusion with regard to recomplementation as a locus of dialectal variation difficult.

The third and final study accounts for the limitations of using offline methods in the previous two studies and the conflation of production and comprehension in study 2. However, the self-paced reading paradigm presents limitations in its own right. While this online task was adopted as a follow-up to Casasanto and Sag (2008), it must be noted that a noncumulative word by word segmentation is an unnatural way to read. Further, natural processing is likely interrupted through forced button clicks. What's more, it may be more ecologically valid to investigate recomplementation through an oral rather than written mode. In present-day Spanish, recomplementation is largely considered a spoken

phenomenon (e.g., Villa-García, 2015). Importantly, the written mode cannot investigate the role of intonation on online comprehension. Spectrographic analysis shows that sentences read aloud, both with and without orthographic commas, include an intonational break between the intervening material and C2 (Villa-García, 2019). Other limitations worthy of mention include assumptions that would benefit from more evidence. We do not have full access to Casasanto and Sag's (2008) statistical report and must make assumptions about nonsignificant findings. Further, in the comparison of our experiment with Casasanto and Sag's, we assume a unified Spanish and English account of recomplementation (Villa-García, 2019). This claim would be strengthened by more theoretical evidence. In any case, we do not offer a direct comparison between studies, as we investigate a different language, a different sentence type, and include a secondary critical region of analysis pertaining to matrix verb biases.

5.3 FUTURE RESEARCH

This dissertation offers several experimental and theoretical avenues for follow-up study. These areas include extending the research to different populations, new experimental methods, as well as isolating and testing individual hypotheses. For example, study 1 is designed to shed light on the divergent outcomes of heritage speaker populations. However, it is not designed to tease apart avoidance of silent phenomena, input as a trigger, and reduction of processing complexity hypotheses. Future experiments on the avoidance of silent elements should include multiple linguistic phenomena that are optionally spelled out (e.g., resumptive pronouns) within the same experiment. The input as a trigger hypothesis can be further investigated by adopting a bilingual comparison group. Ideally this group would represent the input that heritage speakers were exposed to during early

language development. These speaking partners might include immediate and extended family, neighbors and other community members. The processing complexity hypothesis of heritage language development can be more directly tested via online methods. This avenue of research is of particular interest, given growing interest in the relationship between bilingual experience and cognitive and linguistic processing (e.g., Tabori et al., 2018). By comparing heritage speakers to adult second language learners, one is able to consider the role of age of onset of bilingualism and general language experience on processing. This line of research can have an impact on both theoretical and applied literature.

Study 3 in this dissertation offers only the second investigation of recomplementation via online methodology. Specifically, this dissertation and Casasanto and Sag (2008) adopted a self-paced reading paradigm. Given that a memory-based and an expectation-based processing account have been proposed, future studies should continue to test these models. As argued in Staub (2010), these accounts need not be mutually exclusive. Importantly, experiments that adopt different online methods will provide valuable insights. Eye-tracking methods, for example, can offer a finer-grained analysis with multiple dependent measures that more directly tease apart forward-looking and backward-looking behaviors. Not only can they measure reading times at a critical region through eye-fixation duration, but they can also measure regressions (i.e., backward-directed eye movements). Evidence of regression to C1 could serve as strong evidence that C2 is a reiteration of C1 that must be retrieved during complement integration. What's more, given that recomplementation is primarily considered an oral phenomenon in present-day Spanish, methods that facilitate listening as opposed to reading modes offer an even more ecologically valid approach. Neuropsychological methods that measure brain activity either directly (e.g., event related potential, ERP) or indirectly (e.g., functional

magnetic resonance imaging, fMRI) may be superior to behavioral methods for this purpose (e.g., Abutalebi & Della Rosa, 2008). Coincidentally, investigating online recomplementation via listening prompt is also desirable when studying bilingual populations who may not have been formally educated in the language of testing (e.g., Bowles, 2011; Montrul, Foote, & Perpiñán, 2008; Potowski, Jegerski, & Morgan-Short, 2009).

A third avenue of future research is more theoretical in nature. Study 3 proposes the *C2 complexity correlate*, where conventional patterns of C2 lexicalization can be probabilistically predicted along a length of intervener or complexity constraint. This proposal is central to our expectation-based model and is grounded in selected theoretical, experimental and old Spanish corpus-based evidence (e.g., Casasanto & Sag, 2008; Echeverría & López Seoane, 2019; Ledgeway, 2000; Radford, 2018; Villa-García, 2019). However, there is no present-day corpus of recomplementation in Spanish with a representative sample of construction types and written and spoken modes. Such a corpus would provide important evidence in favor of or against the claim that C2 *can* and *should* be predicted as a function of complexity. Lastly, the concept of complexity would benefit from a more precise definition. Even defining complexity as the distance between C1 and C2 is unfortunately vague. As noted by Gibson and colleagues, distance can be measured by number of letters, syllables, words or constituents. Factors other than distance can also theoretically contribute to C2 lexicalization. For example, Echeverría and López Seoane (2019) found that mood, namely the subjunctive as opposed to the indicative, was a significant predictor of C2 lexicalization in old Spanish. Further, in study 1 of this dissertation, we found that heritage speakers' proportion of C2 lexicalization preference is significantly higher in question as opposed to statement items. In summary, much work

remains to further define the C2 complexity correlate, along with potential factors that probabilistically predict C2 lexicalization.

5.4 FINAL REMARKS

Through the window of recomplementation literature, this dissertation has offered evidence in support of three broad claims: (1) speakers with diverse profiles inform general theory; (2) researcher selection bias must not be overlooked; and (3) experimental models should be aligned with syntactic-theoretical accounts. Specifically, bilingualism literature contributes naturally to existing discussions on the role of input and experience in language development, processing complexity, and variation among the Spanishes of the world. Concerns around selection bias in terms of participant and experimental item selection is relevant for all methods of data collection. Lastly, in order to promote comprehensive and interdisciplinary research agendas, experimental and syntactic-theoretical accounts should be in constant conversation with one another and develop in unison. We don't consider any of these claims to be particularly controversial. However, general agreement does not imply shared practices and we have argued that these topics represent limitations that are all too common in the field. Thus, we have found it worthwhile to reiterate the importance of these claims and to offer experimental evidence in their support. In so doing, extending recomplementation studies to bilingual populations, understudied dialects, and online processing complement the robust extant literature in the syntactic-theoretical domain.

Appendices

Appendix A: Linguistic Questionnaire and Self-Assessment

- Sex: Male Female
- Age: _____
- Country of Birth: _____
- If not US born:
 - Age of arrival in the USA _____
 - Length of residence in the USA _____
- Occupation: _____
- What is your first language? _____
- What is the first language of: your mother? _____ your father? _____

- Highest Level of Schooling:
 - Primary High School Community College/Professional University
- Highest Level of Schooling of your mother:
 - Primary High School Community College/Professional University
- Highest Level of Schooling of your father:
 - Primary High School Community College/Professional University

- Did you learn your first language from birth? Yes No
- Which language(s) did you speak at home as a child? _____
- What language do you feel most comfortable with at this time? Spanish / English / Both
- Which language(s) and in what country were you formally educated in?
 - Primary/Elementary School _____
 - High School _____
 - College _____

- Present contact with Spanish & English

At school:							
	Only English	Mainly English	A little more English	Both equally	A little more Spanish	Mainly Spanish	Only Spanish
At home:							
	Only English	Mainly English	A little more English	Both equally	A little more Spanish	Mainly Spanish	Only Spanish
At work:							
	Only English	Mainly English	A little more English	Both equally	A little more Spanish	Mainly Spanish	Only Spanish
In social situations:							
	Only English	Mainly English	A little more English	Both equally	A little more Spanish	Mainly Spanish	Only Spanish

- How often do you visit Spanish speaking countries?

- Frequently Not very frequently Rarely Never

¿For how long? _____

- ¿How often do you watch television in Spanish?

- Frequently Not very frequently Rarely Never

¿For how long? _____

- If you have taken formal language instruction courses in Spanish during elementary school, high school, or college, Approximately how many hours per week at each level?

- Do you currently take formal Spanish courses at the University level? Yes / No

If yes...

- How many hours per week? _____
- For how many consecutive years? _____

Self-Assessment: Please rate your linguistic ability by selecting marks from 0 “basic” to 4 “excellent”.

	1 (basic)	2 (adequate)	3 (good)	4 (excellent)
READING				
English				
Spanish				
Other ()				
WRITING				
English				
Spanish				
Other ()				
SPEAKING				
English				
Spanish				
Other ()				
LISTENING				
English				
Spanish				
Other ()				

Cuestionario Lingüístico y Auto-Evaluación (Spanish Version)

- Sexo: Masculino Femenino
- Edad: _____
- País de nacimiento: _____
 - Si no fue nacido en Colombia/Cuba, ¿Cuántos años tenía cuando llegó? _____
- Ocupación: _____
- ¿Cuál es su lengua nativa? _____
- ¿Cuál es la lengua nativa de su madre? _____ ¿de su padre? _____

- Nivel de escolaridad más alto:
 - Enseñanza media Técnica/Profesional Universitaria
- Nivel de escolaridad más alto de su madre:
 - Enseñanza media Técnica/Profesional Universitaria
- Nivel de escolaridad más alto de su padre:
 - Enseñanza media Técnica/Profesional Universitaria

- ¿Aprendió su lengua nativa desde la infancia? Sí / No
- ¿Qué lengua(s) hablaba usted en la casa de niño? _____
- ¿En qué lengua se siente más cómodo en estos momentos? Español / Inglés / Ambos
- ¿En qué lengua(s) se educó usted formalmente? ¿Y en qué país?
 - Escuela primaria o elemental _____
 - Enseñanza media (preuniversitario) _____
 - Instituto técnico profesional (college) _____
 - Universidad _____

- Contacto actual con el inglés y el español:

En la escuela:							
	Inglés solamente	Mayormente en inglés	Un poco más en inglés	Igual en los dos	Un poco más en español	Mayormente en español	Español solamente
En la casa:							
	Inglés solamente	Mayormente en inglés	Un poco más en inglés	Igual en los dos	Un poco más en español	Mayormente en español	Español solamente
En el trabajo:							
	Inglés solamente	Mayormente en inglés	Un poco más en inglés	Igual en los dos	Un poco más en español	Mayormente en español	Español solamente
En situaciones sociales							
	Inglés solamente	Mayormente en inglés	Un poco más en inglés	Igual en los dos	Un poco más en español	Mayormente en español	Español solamente

- ¿Cuán a menudo visita usted países de habla-inglesa?

Frecuentemente No muy frecuente Casi nunca Nunca

¿Por cuánto tiempo? _____

- ¿Con cuánta frecuencia vea la televisión en inglés?

Frecuentemente No muy frecuente Casi nunca Nunca

¿Por cuánto tiempo? _____

- Si tomó cursos en la instrucción formal del inglés durante la primaria, o escuela secundaria, ¿aproximadamente cuántas horas tomó cada semana? _____

- ¿Sigue tomando estos cursos del inglés formal en la universidad? Sí / No

- ¿Cuántos años ha tomado de estudios consecutivos en inglés? _____

- ¿Cuántas horas de instrucción formal en inglés recibe cada semana? _____

Auto-evaluación: Por favor evalúe su nivel idiomático en cada una de las lenguas que habla dentro de las áreas siguientes:

	1 (básico)	2 (adecuado)	3 (bien)	4 (excelente)
LECTURA				
inglés				
español				
otro ()				
ESCRITURA				
inglés				
español				
otro ()				
EXPRESIÓN ORAL				
inglés				
español				
otro ()				
COMPRENSIÓN				
inglés				
español				
otro ()				

Appendix B: Written Spanish Proficiency Test (DELE)

Parte A: *Escoja la mejor respuesta entre las cuatro opciones para completar las siguientes frases.*

<p>1. Al oír del accidente de su buen amigo, Paco se puso _____ .</p> <p>a. alegre b. fatigado c. hambriento d. desconsolado</p>	<p>2. No puedo comprarlo porque me _____ dinero.</p> <p>a. falta b. dan c. presta d. regalan</p>
<p>3. Tuvo que guardar cama por estar _____ .</p> <p>a. enfermo b. vestido c. ocupado d. parado</p>	<p>4. Aquí está tu café, Juanito. No te quemes, que está muy _____ .</p> <p>a. dulce b. amargo c. agrio d. caliente</p>
<p>5. Al romper los anteojos, Juan se asustó porque no podía _____ sin ellos.</p> <p>a. discurrir b. oír c. ver d. entender</p>	<p>6. ¡Pobrecita! Está resfriada y no puede _____ .</p> <p>a. salir de casa b. recibir cartas c. respirar con pena d. leer las noticias</p>
<p>7. Era una noche oscura sin _____ .</p> <p>a. estrellas b. camas c. lágrimas d. nubes</p>	<p>8. Cuando don Carlos salió de su casa, saludó a un amigo suyo: - Buenos días, _____ .</p> <p>a. ¿Qué va? b. ¿Cómo es? c. ¿Quién es? d. ¿Qué tal?</p>
<p>9. ¡Qué ruido había con los gritos de los niños y el _____ de los perros!</p> <p>a. olor b. sueño c. hambre d. ladrar</p>	<p>10. Para saber la hora, don Juan miró el _____ .</p> <p>a. calendario b. bolsillo c. estante d. despertador</p>

<p>11. Yo, que comprendo poco de mecánica, sé que el auto no puede funcionar sin _____ .</p> <p>a. permiso b. comer c. aceite d. bocina</p>	<p>12. Nos dijo mamá que era hora de comer y por eso _____ .</p> <p>a. fuimos a nadar b. tomamos asiento c. comenzamos a fumar d. nos acostamos pronto</p>
<p>13. ¡Cuidado con ese cuchillo o vas a _____ el dedo!</p> <p>a. cortarte b. torcerte c. comerte d. quemarte</p>	<p>14. Tuvo tanto miedo de caerse que se negó a _____ con nosotros.</p> <p>a. almorzar b. charlar c. cantar d. patinar</p>
<p>15. Abrió la ventana y miró: en efecto, grandes lenguas de _____ salían llameando de las casas.</p> <p>a. zorros b. serpientes c. cuero d. fuego</p>	<p>16. Compró ejemplares de todos los diarios pero en vano. No halló _____ .</p> <p>a. los diez centavos b. el periódico perdido c. la noticia que deseaba d. los ejemplos</p>
<p>17. Por varias semanas acudieron colegas del difunto profesor a _____ el dolor de la viuda.</p> <p>a. aliviar b. dulcificar c. embromar d. estorbar</p>	<p>18. Sus amigos pudieron haberlo salvado pero lo dejaron _____ .</p> <p>a. ganar b. parecer c. perecer d. acabar</p>
<p>19. Al salir de la misa me sentía tan caritativo que no pude menos que _____ a un pobre mendigo que había allí sentado.</p> <p>a. pegarle b. darle una limosna c. echar una mirada d. maldecir</p>	<p>20. Al lado de la Plaza de Armas había dos limosneros pidiendo _____ .</p> <p>a. pedazos b. paz c. monedas d. escopetas</p>

<p>21. Siempre maltratado por los niños, el perro no podía acostumbrarse a _____ de sus nuevos amos.</p> <p>a. las caricias b. los engaños c. las locuras d. los golpes</p>	<p>22. ¿Dónde estará mi cartera? La dejé aquí mismo hace poco y parece que el necio de mi hermano ha vuelto a _____ .</p> <p>a. dejármela b. deshacérmela c. escondérmela d. acabármela</p>
<p>23. Permaneció un gran rato abstraído, los ojos clavados en el fogón y el pensamiento _____ .</p> <p>a. en el bolsillo b. en el fuego c. lleno de alboroto d. Dios sabe dónde</p>	<p>24. En vez de dirigir el tráfico estabas charlando, así que tú mismo _____ del choque.</p> <p>a. sabes la gravedad b. eres testigo c. tuviste la culpa d. conociste a las víctimas</p>
<p>25. Posee esta tierra un clima tan propio para la agricultura como para _____ .</p> <p>a. la construcción de trampas b. el fomento de motines c. el costo de vida d. la cría de reses</p>	<p>26. Aficionado leal de obras teatrales, Juan se entristeció al saber _____ del gran actor.</p> <p>a. del fallecimiento b. del éxito c. de la buena suerte d. de la alabanza</p>
<p>27. Se reunieron a menudo para efectuar un tratado pero no pudieron _____ .</p> <p>a. desavenirse b. echarlo a un lado c. rechazarlo d. llevarlo a cabo.</p>	<p>28. Se negaron a embarcarse porque tenían miedo de _____ .</p> <p>a. los peces b. los naufragios c. los faros d. las playas</p>
<p>29. La mujer no aprobó el cambio de domicilio pues no le gustaba _____ .</p> <p>a. el callejeo b. el puente c. esa estación d. aquel barrio</p>	<p>30. Era el único que tenía algo que comer pero se negó a _____ .</p> <p>a. hojearlo b. ponérselo c. conservarlo d. repartirlo</p>

Parte B: *En la siguiente lectura hay unos espacios en blanco (1-20). Lea la lectura y después llena los espacios con la mejor respuesta según la lista de tres opciones que encuentras en la próxima página, la hoja de respuestas.*

El sueño de Juan Miró

Hoy se inaugura en Palma de Mallorca la Fundación Pilar y Joan Miró, en el mismo lugar en donde el artista vivió sus últimos treinta y cinco años. El sueño de Joan Miró se ha _____ (1). Los fondos donados a la ciudad por el pintor y su esposa en 1981 permitieron que el sueño se _____ (2); más tarde, en 1986, el Ayuntamiento de Palma de Mallorca decidió _____ (3) al arquitecto Rafael Moneo un edificio que _____ (4) a la vez como sede de la entidad y como museo moderno. El proyecto ha tenido que _____ (5) múltiples obstáculos de carácter administrativo. Miró, coincidiendo _____ (6) los deseos de toda su familia, quiso que su obra no quedara expuesta en ampulosos panteones de arte o en _____ (7) de coleccionistas acaudalados; por ello, en 1981, creó la fundación mallorquina. Y cuando estaba _____ (8) punto de morir, donó terrenos y edificios, así como las obras de arte que en ellos _____ (9).

El edificio que ha construido Rafael Moneo se enmarca en _____ (10) se denomina "Territorio Miró", espacio en el que se han _____ (11) de situar los distintos edificios que constituyen la herencia del pintor.

El acceso a los mismos quedará _____ (12) para evitar el deterioro de las obras. Por otra parte, se _____ (13), en los talleres de grabado y litografía, cursos _____ (14) las distintas técnicas de estampación. Estos talleres también se cederán periódicamente a distintos artistas contemporáneos, _____ (15) se busca que el "Territorio Miró" _____ (16) un centro vivo de creación y difusión del arte a todos los _____ (17).

La entrada costará 500 pesetas y las previsiones dadas a conocer ayer aspiran _____ (18) que el centro acoja a unos 150.000 visitantes al año. Los responsables esperan que la institución funcione a _____ (19) rendimiento a principios de la _____ (20) semana, si bien el catálogo completo de las obras de la Fundación Pilar y Joan Miró no estará listo hasta dentro de dos años.

Hoja de respuestas

- | | | |
|---|---|--|
| 1. a. cumplido
b. completado
c. terminado | 2. a. inició
b. iniciara
c. iniciaba | 3. a. encargar
b. pedir
c. mandar |
| 4. a. hubiera servido
b. haya servido
c. sirviera | 5. a. superar
b. enfrentarse
c. acabar | 6. a. por
b. en
c. con |
| 7. a. voluntad
b. poder
c. favor | 8. a. al
b. en
c. a | 9. a. habría
b. había
c. hubo |
| 10. a. que
b. el que
c. lo que | 11. a. pretendido
b. tratado
c. intentado | 12. a. disminuido
b. escaso
c. restringido |
| 13. a. darán
b. enseñarán
c. dirán | 14. a. sobre
b. en
c. para | 15. a. ya que
b. así
c. para |
| 16. a. será
b. sea
c. es | 17. a. casos
b. aspectos
c. niveles | 18. a. a
b. de
c. para |
| 19. a. total
b. pleno
c. entero | 20. a. siguiente
b. próxima
c. pasada | |

Appendix C: Stimuli for Acceptability Judgment Task

Item	Sentence Type	C2	Stimuli
1	statement	que	<p>Preamble: Esa casita antigua, voy a pintarla. “I will paint the old house.”</p> <p>Response: Me dijo que esa casita antigua que iba a pintarla. “S/he told me s/he was going to paint the old house.”</p>
2	statement	que	<p>Preamble: Esas joyas elegantes, voy a llevarlas. “I will wear the elegant jewels.”</p> <p>Response: Me dijo que esas joyas elegantes que iba a llevarlas. “S/he told me s/he was going to wear the elegant jewels.”</p>
3	statement	que	<p>Preamble: Esa motocicleta clásica, voy a montarla. “I will ride the classic motorcycle.”</p> <p>Response: Me dijo que esa motocicleta clásica que iba a montarla. “S/he told me s/he was going to ride the classic motorcycle.”</p>
4	statement	que	<p>Preamble: Esa guitarra eléctrica, voy a venderla. “I will sell the electric guitar.”</p> <p>Response: Me dijo que esa guitarra eléctrica que iba a venderla. “S/he told me s/he was going to sell the electric guitar.”</p>
5	statement	que	<p>Preamble: Ese coche deportivo, voy a comprarlo. “I will buy the sports car.”</p> <p>Response: Me dijo que ese coche deportivo que iba a comprarlo. “S/he told me s/he was going to buy the sports car.”</p>
6	statement	que	<p>Preamble: Esa camisa rota, voy a coserla. “I will sew the torn shirt.”</p> <p>Response: Me dijo que esa camisa rota que iba a coserla. “S/he told me s/he was going to sew the torn shirt.”</p>
7	statement	no que	<p>Preamble: Ese traje formal, voy a pedirlo.</p>

			<p>“I will order the formal suit.”</p> <p>Response: Me dijo que ese traje formal iba a pedirlo. “S/he told me s/he was going to order the formal suit.”</p>
8	statement	no que	<p>Preamble: Ese folleto informativo, voy a distribuirlo. “I will distribute the informational flier.”</p> <p>Response: Me dijo que ese folleto informativo iba a distribuirlo. “S/he told me s/he was going to distribute the informational flier.”</p>
9	statement	no que	<p>Preamble: Ese libro clásico, voy a leerlo. “I will read the classic book.”</p> <p>Response: Me dijo que ese libro clásico iba a leerlo. “S/he told me s/he was going to read the classic book.”</p>
10	statement	no que	<p>Preamble: Esa canción popular, voy a buscarla. “I will search for the popular song.”</p> <p>Response: Me dijo que esa canción popular iba a buscarla. “S/he told me s/he was going to search for the popular song.”</p>
11	statement	no que	<p>Preamble: Esa clase nocturna, voy a tomarla. “I will take the night class.”</p> <p>Response: Me dijo que esa clase nocturna iba a tomarla. “S/he told me s/he was going to take the night class.”</p>
12	statement	no que	<p>Preamble: Esa planta seca, voy a regarla. “I will water the dry plant.”</p> <p>Response: Me dijo que esa planta seca iba a regarla. “S/he told me s/he was going to water the dry plant.”</p>
13	question	que	<p>Preamble: Ese postre dulce, ¿dónde vas a guardarlo? “Where will you store the sweet dessert?”</p> <p>Response: Me dijo que ese postre dulce que dónde iba a guardarlo. “S/he asked me where I was going to store the sweet dessert.”</p>

14	question	que	<p>Preamble: Ese teléfono viejo, ¿cuándo vas a cambiarlo? “When will you change the old telephone?”</p> <p>Response: Me dijo que ese teléfono viejo que cuándo iba a cambiarlo. “S/he asked me when I was going to change the old telephone.”</p>
15	question	que	<p>Preamble: Ese pescado frito, ¿cuándo vas a cocinarlo? “When will you cook the fried fish?”</p> <p>Response: Me dijo que ese pescado frito que cuándo iba a cocinarlo. “S/he asked me when I was going to cook the fried fish.”</p>
16	question	que	<p>Preamble: Esa bicicleta nueva, ¿cuándo vas a montarla? “When will you ride the new bicycle?”</p> <p>Response: Me dijo que esa bicicleta nueva que cuándo iba a montarla. “S/he asked me when I was going to ride the new bicycle.”</p>
17	question	que	<p>Preamble: Ese uniforme colombiano, ¿dónde vas a encontrarlo? “Where will you find the Colombian uniform?”</p> <p>Response: Me dijo que ese uniforme colombiano que dónde iba a encontrarlo. “S/he asked me where I was going to find the Colombian uniform.”</p>
18	question	que	<p>Preamble: Esa camisa fea, ¿cuándo vas a devolverla? “When will you return the ugly shirt?”</p> <p>Response: Me dijo que esa camisa fea que cuándo iba a devolverla. “S/he asked me when I was going to return the ugly shirt.”</p>
19	question	no que	<p>Preamble: Esa flor morada, ¿dónde vas a sembrarla? “Where will you plant the purple flower?”</p> <p>Response: Me dijo que esa flor morada dónde iba a sembrarla. “S/he asked me where I was going to plant the purple flower.”</p>

20	question	no que	<p>Preamble: Ese dibujo bonito, ¿dónde vas a colgarlo? “Where will you hang the beautiful drawing?”</p> <p>Response: Me dijo que ese dibujo bonito dónde iba a colgarlo. “S/he asked me where I was going to hang the beautiful drawing.”</p>
21	question	no que	<p>Preamble: Ese dinero estadounidense, ¿dónde vas a cambiarlo? “Where will you exchange the US currency?”</p> <p>Response: Me dijo que ese dinero estadounidense dónde iba a cambiarlo. “S/he asked me where I was going to exchange the US currency.”</p>
22	question	no que	<p>Preamble: Esa mesa pesada, ¿cómo vas a moverla? “How will you move the heavy table?”</p> <p>Response: Me dijo que esa mesa pesada cómo iba a moverla. “S/he asked me how I was going to move the heavy table.”</p>
23	question	no que	<p>Preamble: Esa chaqueta roja, ¿cuándo vas a comprarla? “When will you buy the red jacket?”</p> <p>Response: Me dijo que esa chaqueta roja cuándo iba a comprarla. “S/he asked me when I was going to buy the red jacket.”</p>
24	question	no que	<p>Preamble: Ese museo privado, ¿cuándo vas a visitarlo? “When will you visit the private museum?”</p> <p>Response: Me dijo que ese museo privado cuándo iba a visitarlo. “S/he asked me when I was going to visit the private museum.”</p>

Appendix D: Stimuli for Preference Task

Item	Sentence Type	Stimuli
1	Statement	<p>Preamble: Ayer Leonardo tuvo que recordarme del folleto que creamos la semana pasada. “Yesterday Leo reminded me about the flyer that we created last week.”</p> <p>Option 1: Leonardo me dijo que ese folleto, iba a distribuirlo en el centro. Option 2: Leonardo me dijo que ese folleto, que iba a distribuirlo en el centro. “Leo told me that he was going to distribute the flyer downtown.”</p>
2	Statement	<p>Preamble: Ayer tuve que recordarle a Natalia de los conciertos que el músico iba a presentar esta semana. “Yesterday I reminded Natalie of the concerts that the musician was going to present this week.”</p> <p>Option 1: Yo le dije que ese concierto, iba a asistirlo este viernes. Option 2: Yo le dije que ese concierto, que iba a asistirlo este viernes. “I told him that I was going to attend the concert this Friday.”</p>
3	Statement	<p>Preamble: Ayer Pablo tuvo que recordarme de la opción de alquilar la computadora de la biblioteca. “Yesterday Pablo reminded me of the option of renting a computer from the library.”</p> <p>Option 1: Pablo me dijo que esa computadora, que iba a alquilarla toda la semana. Option 2: Pablo me dijo que esa computadora, iba a alquilarla toda la semana. “Pablo told me that he was going to rent the computer for the entire week.”</p>
4	Statement	<p>Preamble: Ayer Raúl tuvo que recordarme de su sombrero que no había llevado por mucho tiempo. “Yesterday Raul reminded me about his hat which he hadn’t worn for a while.”</p> <p>Option 1: Raúl me dijo que ese sombrero, iba a llevarlo por la</p>

		<p>tarde.</p> <p>Option 2: Raúl me dijo que ese sombrero, que iba a llevarlo por la tarde.</p> <p>“Raul told me that he was going to wear the hat in the afternoon.”</p>
5	Statement	<p>Preamble: Ayer, como la semana pasada, Miguel tuvo que recordarme del anillo que había visto en la joyería.</p> <p>“Yesterday, like last week, Miguel reminded me of the ring that he had seen in the jewelry store.”</p> <p>Option 1: Miguel me repitió que ese anillo, que iba a comprarlo un día pronto.</p> <p>Option 2: Miguel me repitió que ese anillo, iba a comprarlo un día pronto.</p> <p>“Miguel told me again that he was going to buy the ring one day soon.”</p>
6	Statement	<p>Preamble: Ayer, como la semana pasada, tuve que recordarle a Alfredo de lo que iba a hacer con la camisa fea.</p> <p>“Yesterday, like last week, I reminded Alfredo what I was going to do with the ugly shirt”</p> <p>Option 1: Yo le repetí que esa camisa, que iba a llevarla al cumpleaños.</p> <p>Option 2: Yo le repetí que esa camisa, iba a llevarla al cumpleaños.</p> <p>“I told him again that I was going to wear the shirt for the birthday party.”</p>
7	Statement	<p>Preamble: Ayer, como la semana pasada, tuve que recordarle a Javier de la hora que iba a tomar la clase.</p> <p>“Yesterday, like last week, I had to remind Javier of the time that I was going to take the class.”</p> <p>Option 1: Yo le repetí que esa clase, iba a tomarla por la tarde.</p> <p>Option 2: Yo le repetí que esa clase, que iba a tomarla por la tarde.</p> <p>“I told him again that I was going to take the class in the afternoon.”</p>
8	Statement	<p>Preamble: Ayer, como la semana pasado, tuve que recordarle a Ramón de la cama.</p> <p>“Yesterday, like last week, I reminded Ramon about the bed.”</p> <p>Option 1: Yo le repetí que esa cama, iba a comprarla pronto.</p>

		<p>Option 2: Yo le repetí que esa cama, que iba a comprarla pronto. “I told him again that I was going to buy the bed soon.”</p>
9	question	<p>Preamble: Ayer, Felipe tuvo que recordarme de la renovación de la casa. “Yesterday, Philip reminded me about the home renovation.”</p> <p>Option 1: Felipe me preguntó que esa casa, que cuándo iba a renovarla. Option 2: Felipe me preguntó que esa casa, cuándo iba a renovarla. “Philip asked me when I was going to renovate the home.”</p>
10	question	<p>Preamble: Ayer, tuve que recordarle a Mario de la colección de joyería. “Yesterday I reminded Mario of the jewelry collection.”</p> <p>Option 1: Yo le pregunté que esa joyería, adónde iba a llevarla. Option 2: Yo le pregunté que esa joyería, que adónde iba a llevarla. “I asked him where he was going to take the jewelry.”</p>
11	question	<p>Preamble: Ayer, como la semana pasada, tuve que recordarle a Ramón del teléfono antiguo. “Yesterday, like last week, I reminded Ramon about the old telephone.”</p> <p>Option 1: Yo le pregunté que ese teléfono, que cuándo iba a cambiarlo. Option 2: Yo le pregunté que ese teléfono, cuándo iba a cambiarlo. “I asked him when he was going to exchange the telephone.”</p>
12	question	<p>Preamble: Ayer, como la semana pasada, Rodrigo tuvo que recordarme del tamaño del árbol. “Yesterday, like last week, Rodrigo reminded me of the size of the tree.”</p> <p>Option 1: Rodrigo me preguntó que ese árbol, dónde iba a sembrarlo. Option 2: Rodrigo me preguntó que ese árbol, que dónde iba a sembrarlo. “Rodrigo asked me where I was going to plant the tree.”</p>

13	question	<p>Preamble: Ayer, Carlos tuvo que recordarme que no íbamos a dejar el postre en la mesa. “Yesterday, Carlos reminded me that we weren’t going to leave the dessert on the table.”</p> <p>Option 1: Carlos me dijo que ese postre, que adónde iba a guardarlo.</p> <p>Option 2: Carlos me dijo que ese postre, adónde iba a guardarlo. “Carlos asked me where I was going to leave the dessert.”</p>
14	question	<p>Preamble: Ayer, como la semana pasada, tuve que recordarle a Emilia de la bicicleta en nuestro garaje. “Yesterday, like last week, I reminded Emilia about the bicycle in our garage.”</p> <p>Option 1: Yo le dije que esa bicicleta, que cuándo iba a montarla.</p> <p>Option 2: Yo le dije que esa bicicleta, cuándo iba a montarla. “I asked her when she was going to ride the bicycle.”</p>
15	question	<p>Preamble: Ayer, como la semana pasada, Miguel tuvo que recordarme de la chaqueta que habíamos visto. “Yesterday, like last week, Miguel reminded me about the jacket that we had seen.”</p> <p>Option 1: Miguel me repitió que esa chaqueta, cuándo iba a comprarla.</p> <p>Option 2: Miguel me repitió que esa chaqueta, que cuándo iba a comprarla. “Miguel asked me again when I was going to buy the jacket.”</p>
16	question	<p>Preamble: Ayer, como la semana pasada, tuve que recordarle a María del dibujo en el suelo. “Yesterday, like last week, I reminded Maria of the picture on the floor.”</p> <p>Option 1: Yo le repetí que ese dibujo, dónde iba a colgarlo.</p> <p>Option 2: Yo le repetí que ese dibujo, que dónde iba a colgarlo. “I asked her again where she was going to hang the painting.”</p>

Appendix E: Stimuli for Elicited Imitation Task

Item	Sentence Type	Intervener Type	C2 + Length	Sentence
1	statement	indirect object	que_short	Me dice que al profesor que le va a enviar una carta de gracias para su jubilación pendiente. “S/he tells me s/he is going to send a thank you card to the professor for his pending retirement.”
2	question	indirect object	que_short	Me pregunta que al doctor que cuándo le voy a dar mi historial médico y pedir los medicamentos. “S/he asks me when I will give the doctor my medical history and request the medications.”
3	statement	direct object	que_short	Me dice que esa camisa que la va a cambiar para una más grande lo más pronto posible. “S/he tells me s/he will exchange the shirt for a bigger one as soon as possible.”
4	question	direct object	que_short	Me pregunta que esa semilla que cuándo la voy a sembrar en el jardín con las otras plantas. “S/he asks me when I will plant the seed in the garden with the other plants.”
5	statement	adverbial	que_short	Me dice que por supuesto que va a apoyar al candidato a gobernador de su linda ciudad natal. “S/he tells me s/he will of course support the candidate for governor from her wonderful city of birth.”
6	question	adverbial	que_short	Me pregunta que esta noche que dónde voy a querer salir a comer como es mi turno escoger. “S/he asks me where I will want to go out to eat tonight since it is my turn to choose.”
7	statement	indirect object	no que_short	Me dice que al estudioso le va a prestar el libro de práctica antes del examen final.

				“S/he tells me s/he will loan the studious one the practice book before the final exam.”
8	question	indirect object	no que_short	Me pregunta que al deportista cuándo le va a dar el premio prestigioso del jugador más valioso. “S/he asks me when they will give the sportsman the prestigious award for most valuable player.”
9	statement	direct object	no que_short	Me dice que ese museo lo va a visitar en la tarde con unos compañeros de clase. “S/he tells me s/he will visit the museum with a few classmates in the afternoon.”
10	question	direct object	no que_short	Me pregunta que esa canción cuándo la voy a grabar para mi álbum proyectado a salir pronto. “S/he asks me when I will record the song for my album projected to be released soon.”
11	statement	adverbial	no que_short	Me dice que por fin va a graduarse de la universidad con el título de ingeniero civil. “S/he tells me s/he will finally graduate from the university with the title of civil engineer.”
12	question	adverbial	no que_short	Me pregunta que más tarde dónde voy a ir para tomar un descanso después de trabajar tanto. “S/he asks me where I will go later to take a break after working so much.”
13	statement	indirect object	que_long	Me dice que al mesero por haber traducido el menú que le va a dar una propina generosa. “S/he tells me s/he will give the waiter a generous tip for having translated the menu.”
14	question	indirect object	que_long	Me pregunta que al director de administración de la escuela que cuándo le voy a cocinar una cena.

				“S/he asks me when I will cook a dinner for the school’s director of administration.”
15	statement	direct object	que_long	Me dice que esa pintura clásica colgada en su sala que la va a vender la próxima semana. “S/he tells me s/he will sell the classic painting hanging in the living room next week.”
16	question	direct object	que_long	Me pregunta que ese uniforme de la selección de Argentina que cuándo lo voy a pedir como regalo. “S/he asks me when I will request the uniform of the Argentine national team as a gift.”
17	statement	adverbial	que_long	Me dice que por suerte después de conseguir la entrada que va a tener suficiente tiempo de comer. “S/he tells me s/he will luckily have enough time to eat after obtaining the ticket.”
18	question	adverbial	que_long	Me pregunta que esta tarde después de correr el maratón que dónde voy a celebrar mi logro tremendo. “S/he asks me where I will celebrate my tremendous achievement this afternoon after running in the marathon.”
19	statement	indirect object	no que_long	Me dice que a su compañero del equipo de béisbol le va a vender su nuevo uniforme. “S/he tells me s/he will sell her new uniform to her baseball teammate.”
20	question	indirect object	no que_long	Me pregunta que a mi familia viviendo por todas partes cuándo le voy a compartir las noticias. “S/he asks me when I will share the news with my family living all over the place.”
21	statement	direct object	no que_long	Me dice que esa clase nocturna los martes y jueves la va a tomar con sus amigos.

				“S/he tells me s/he will take the night class on Tuesdays and Thursdays with her friends.”
22	question	direct object	no que_long	Me pregunta que esa película sobre los extraterrestres malos dónde la voy a ver esta tarde. “S/he asks me where I will go this afternoon to see the movie about the evil aliens.”
23	statement	adverbial	no que_long	Me dice que sin duda después de la tormenta severa va a haber mucho daño al techo. “S/he tells me that without a doubt there will be a lot of damage to the roof after the severe storm.”
24	question	adverbial	no que_long	Me pregunta que pasado mañana después de salir del trabajo dónde voy a pasar mi tarde libre. “S/he asks me where I will spend my free afternoon the day after tomorrow after work.”

Appendix F: Stimuli for Sentence Completion Task

Item	Sentence Type	Intervener Type	Length	Stimuli
1	statement	object	short	<p>Preamble: Al paciente, le voy a prescribir mucho descanso. “I will prescribe a lot of rest to the patient.”</p> <p>Prompt and expected response: Me dice que al paciente ... <i>(que) le va a prescribir mucho descanso.</i> “S/he tells me that s/he will prescribe a lot of rest to the patient.”</p>
2	statement	object	short	<p>Preamble: Ese postre, lo voy a dejar en la cocina. “I will leave the dessert in the kitchen.”</p> <p>Prompt and expected response: Me dice que ese postre ... <i>(que) lo va a dejar en la cocina.</i> “S/he tells me that s/he will leave the dessert in the kitchen.”</p>
3	question	object	short	<p>Preamble: Al viejo, ¿cuándo le vas a ofrecer un trabajo? “When will you offer work to the old man?”</p> <p>Prompt and expected response: Me pregunta que al viejo ... <i>(que) cuándo le voy a ofrecer un trabajo.</i> “S/he asks me when I will offer work to the old man.”</p>
4	question	object	short	<p>Preamble: Ese coche, ¿dónde lo vas a dejar para mis padres? “Where will you leave the car for my parents?”</p> <p>Prompt and expected response: Me pregunta que ese coche ... <i>(que) dónde lo voy a dejar para mis padres.</i> “S/he asks me where I will leave the car for my parents.”</p>

5	statement	adverbial	short	<p>Preamble: Más tarde, voy a comprar una almohada cómoda. “Later I will buy a comfortable pillow.”</p> <p>Prompt and expected response: Me dice que más tarde ... (<i>que</i>) va a comprar una almohada cómoda. “S/he tells me that later s/he will buy a comfortable pillow.”</p>
6	statement	adverbial	short	<p>Preamble: Si nieva, voy a volver a casa después del partido. “I will return home after the game if it snows.”</p> <p>Prompt and expected response: Me dice que si nieva ... (<i>que</i>) va a volver a casa después del partido. “S/he tell me that s/he will return home after the game if it snows.”</p>
7	question	adverbial	short	<p>Preamble: Esta noche, ¿dónde vas a salir a comer helado? “Where will you go out for ice cream tonight?”</p> <p>Prompt and expected response: Me pregunta que esta noche ... (<i>que</i>) dónde voy a salir a comer helado. “S/he asks me where I will go out for ice cream tonight.”</p>
8	question	adverbial	short	<p>Preamble: Si escribes, ¿dónde vas a publicar el trabajo? “Where will you publish the work if you write?”</p> <p>Prompt and expected response: Me pregunta que si escribo ... (<i>que</i>) dónde voy a publicar el trabajo. “S/he asks me where I will publish the work if I write.” “He asks me that if I write (that) where I will publish my work after having finished a few essays.”</p>

				<p>Preamble: Al artista reconocido por sus dibujos abstractos, le voy a ofrecer un contrato competitivo. “I will offer a competitive contract to the artist renowned for his abstract drawings.”</p> <p>Prompt and expected response: Me dice que al artista reconocido por sus dibujos abstractos ... <i>(que) le va a ofrecer un contrato competitivo.</i> “S/he tells me s/he will offer a competitive contract to the artist renowned for his abstract drawings.”</p>
9	statement	object	long	
				<p>Preamble: Ese desayuno planeado para el próximo sábado, lo voy a organizar con mi familia. “I will organize with my family the breakfast planned for next Saturday.”</p> <p>Prompt and expected response: Me dice que ese desayuno planeado para el próximo sábado ... <i>(que) lo va a organizar con su familia.</i> “S/he tells me that s/he will organize with her family the breakfast planned for next Saturday.”</p>
10	statement	object	long	
				<p>Preamble: Al estudiante de mi clase de geometría, ¿cuándo le vas a escribir una recomendación? “When will you write a recommendation for the student in your geometry class?”</p> <p>Prompt and expected response: Me pregunta que al estudiante de mi clase de geometría... <i>(que) cuándo le voy a escribir una recomendación.</i> “S/he asks me when I will write a recommendation for the student in my geometry class.”</p>
11	question	object	long	

12	question	object	long	<p>Preamble: Esa maleta para viajes de larga distancia, ¿dónde la vas a llevar esta vez? “Where will you take the suitcase used for long distance trips this time?”</p> <p>Prompt and expected response: Me pregunta que esa maleta para viajes de larga distancia ... <i>(que) dónde la voy a llevar esta vez.</i> “S/he asks me where I will take the suitcase used for long distance trips this time.”</p>
13	statement	adverbial	long	<p>Preamble: Esta mañana después de devolver la blusa, voy a buscar un regalo más barato. “I will look for a cheaper gift this morning after returning the blouse.”</p> <p>Prompt and expected response: Me dice que esta mañana después de devolver la blusa ... <i>(que) va a buscar un regalo más barato.</i> “S/he says s/he will look for a cheaper gift this morning after returning the blouse.”</p>
14	statement	adverbial	long	<p>Preamble: Si llega al festival con suficiente tiempo, voy a asistir al baile de salsa. “I will attend the salsa dance if I arrive at the festival with enough time.”</p> <p>Prompt and expected response: Me dice que si llega al festival con suficiente tiempo ... <i>(que) va a asistir al baile de salsa.</i> “S/he says s/he will attend the salsa dance if s/he arrives at the festival with enough time.”</p>
15	question	adverbial	long	<p>Preamble: Esta tarde durante la lección de violín, ¿cómo le vas a agradecer al instructor? “How will you thank the instructor during the violin lesson this afternoon?”</p> <p>Prompt and expected response: Me pregunta que esta tarde durante la lección de violín... <i>(que) cómo voy a agradecerle al instructor.</i></p>

				“S/he asks me how I will thank the instructor during the violin lesson this afternoon.”
16	question	adverbial	long	<p>Preamble: Si ganas la competencia de arte anual, ¿cómo vas a gastar el premio monetario? “How will you spend the monetary prize if you win the annual art competition?”</p> <p>Prompt and expected response: Me pregunta que si gano la competencia de arte anual ... <i>(que) cómo voy a gastar el premio monetario.</i> “S/he asks me how I will spend the monetary prize if I win the annual art competition.”</p>

Appendix G: Sample Stimuli for Reverse Digit Span Task

Trial	Stimuli	Correct Response
1	7 8	8 7
2	3 5	5 3
3	8 5 2	2 5 8
4	3 9 5	5 9 3
5	8 4 2 3	3 2 4 8
6	8 7 9 1	1 9 7 8
7	3 1 8 3 4	4 3 8 1 3
8	3 8 1 8 9	9 8 1 8 3
9	2 2 6 2 5 4	4 5 2 6 2 2
10	3 8 8 9 3 6	6 3 9 8 8 3
11	4 9 5 2 4 7 3	3 7 4 2 5 9 4
12	1 4 6 3 4 9 6	6 9 4 3 6 4 1
13	4 9 3 8 8 1 5 8	8 5 1 8 8 3 9 4
14	1 8 7 2 1 5 5 8	8 5 5 1 2 7 8 1
15	8 2 5 3 7 1 9 1 6	6 1 9 1 7 3 5 2 8
16	6 2 8 2 9 7 5 4 9	9 4 5 7 9 2 8 2 6

Appendix H: Master List for Self-Paced Reading Task

Item	Sentence
1	<p>Él me pregunta/dice que al abogado (de mi defensa) (que) cuándo le envío un correo detallado. "He asks me when I will send my defense lawyer a detailed email."</p>
2	<p>Él me pregunta/dice que al alumno (de mi academia) (que) cuándo le escribo una recomendación fuerte. "He asks me when I will write the student at my academy a strong recommendation."</p>
3	<p>Él me pregunta/dice que al caballo (de la carrera) (que) cuándo le sirvo la comida nutritiva. "He asks me when I will serve the race horse the nutritious food."</p>
4	<p>Él me pregunta/dice que al camarero (de la cantina) (que) cuándo le cuento toda mi historia. "He asks me when I will tell the waiter of the cantina my entire life story."</p>
5	<p>Él me pregunta/dice que al cómico (en el evento) (que) cuándo le ofrezco un contrato anual. "He asks me when I will offer the comic at the event an annual contract."</p>
6	<p>Él me pregunta/dice que al conejo (de mi hermano) (que) cuándo le construyo un cobijo afuera. "He asks me when I will construct an outdoor shelter for my brother's rabbit ."</p>
7	<p>Él me pregunta/dice que al experto (de la materia) (que) cuándo le muestro mi trabajo cumplido. "He asks me when I will show the content expert my finished work."</p>
8	<p>Él me pregunta/dice que al individuo (en la pelea) (que) cuándo le comparto las noticias graves. "He asks me when I will share the serious news with the individual in the fight."</p>
9	<p>Él me pregunta/dice que al médico (de la clínica) (que) cuándo le entrego el cuestionario personal. "He asks me when I will turn in the personal questionnaire to the doctor from the clinic."</p>
10	<p>Él me pregunta/dice que al músico (en el concurso) (que) cuándo le presento el premio prestigioso. "He asks me when I will present the prestigious award to the musician in the competition."</p>

11	<p>Él me pregunta/dice que al negocio (de mi hermano) (que) cuándo le dono mi tiempo valeroso.</p> <p>"He asks me when I will donate my valuable time to my brother's business."</p>
12	<p>Él me pregunta/dice que al sombrero (para el evento) (que) cuándo le añado la pluma gigante.</p> <p>"He asks me when I will add the gigantic feather to the hat for the event."</p>
13	<p>Él me pregunta/dice que esa bicicleta (en el sótano) (que) cuándo la termino para mi hijo.</p> <p>"He asks me when I will finish the bicycle in the basement for my son."</p>
14	<p>Él me pregunta/dice que esa camisa (para mi hermano) (que) cuándo la cambio para la chaqueta.</p> <p>"He asks me when I will exchange the shirt intended for my brother with the jacket."</p>
15	<p>Él me pregunta/dice que esa cocina (en la iglesia) (que) cuándo la reparo para mi comunidad.</p> <p>"He asks me when I will repair the kitchen in the church for the community."</p>
16	<p>Él me pregunta/dice que esa corbata (en el armario) (que) cuándo la lavo para mi hijo.</p> <p>"He asks me when I will wash the tie in the wardrobe for my son."</p>
17	<p>Él me pregunta/dice que esa figura (de la novela) (que) cuándo la pinto en mi cuarto.</p> <p>"He asks me when I will paint in my room that statue from the novel."</p>
18	<p>Él me pregunta/dice que esa película (sobre el futuro) (que) cuándo la muestro en mi fiesta.</p> <p>"He asks me when I will show the futuristic movie at my party."</p>
19	<p>Él me pregunta/dice que esa tarea (en la mochila) (que) cuándo lo completo con mi hermano.</p> <p>"He asks me when I will complete the homework in my backpack with my brother."</p>
20	<p>Él me pregunta/dice que esa tubería (con el óxido) (que) cuándo la arreglo para mi cliente.</p> <p>"He asks me when I will fix the rusty pipes for my client."</p>
21	<p>Él me pregunta/dice que ese boleto (para el teatro) (que) cuándo lo consigo de la vendedora.</p> <p>"He asks me when I will obtain the theatre ticket from the seller."</p>
22	<p>Él me pregunta/dice que ese edificio (con el mercado) (que) dónde lo construyo en el centro.</p> <p>"He asks me where downtown I will construct the building with the market."</p>

23	<p>Él me pregunta/dice que ese museo (en la esquina) (que) cuándo lo visito con mis amigos. "He asks me when I will visit the museum on the corner with my friends."</p>
24	<p>Él me pregunta/dice que ese video (con el pájaro) (que) cuándo lo vendo a la agencia. "He asks me when I will sell the bird video to the agency."</p>
25	<p>Ella me pregunta/dice que al amigo (de mi abuelo) (que) cuándo le regalo la entrada extra. "She asks me when I will give the extra ticket to her grandmother's friend."</p>
26	<p>Ella me pregunta/dice que al carpintero (de la piscina) (que) cuándo le ofrezco una bebida fría. "She asks me when I will offer the pool carpenter a cold refreshment."</p>
27	<p>Ella me pregunta/dice que al compañero (de mi colegio) (que) cuándo le presto el repaso útil. "She asks me when I will lend the helpful review to my classmate."</p>
28	<p>Ella me pregunta/dice que al empleado (de mi compañía) (que) cuándo le compro una computadora nueva. "She asks me when I will buy a new computer for the employee at my company."</p>
29	<p>Ella me pregunta/dice que al instituto (en mi vecindario) (que) cuándo le envío el documento oficial. "She asks me when I will send the official document to the institute in my neighborhood."</p>
30	<p>Ella me pregunta/dice que al monumento (de la tragedia) (que) cuándo le añado unas flores coloridas. "She asks me when I will add colorful flowers to the monument from the tragedy."</p>
31	<p>Ella me pregunta/dice que al movimiento (de la academia) (que) cuándo le dono mi apoyo financiero. "She asks me when I will donate my financial support to the university movement."</p>
32	<p>Ella me pregunta/dice que al muchacho (con el dinero) (que) cuándo le muestro la joyería elegante. "She asks me when I will show the elegant jewelry to the wealthy teenager."</p>
33	<p>Ella me pregunta/dice que al perrito (de mi vecino) (que) cuándo le sirvo la nueva comida. "She asks me when I will serve the new food to my neighbor's dog."</p>
34	<p>Ella me pregunta/dice que al secretario (de mi consultorio) (que) cuándo le ofrezco un descanso merecido.</p>

	"She asks me when I will offer a well-deserved break to the office secretary."
35	Ella me pregunta/dice que al sustituto (de la escuela) (que) cuándo le ofrezco un trabajo sustancial. "She asks me when I will offer substantial work to the substitute at the school."
36	Ella me pregunta/dice que al voluntario (en el refugio) (que) cuándo le cocino una cena elaborada. "She asks me when I will cook an elaborate dinner for the shelter volunteer."
37	Ella me pregunta/dice que esa autora (de la novela) (que) dónde la encuentro en la mañana. "She asks me where I will find author of the novel in the morning."
38	Ella me pregunta/dice que esa casita (de mi abuelo) (que) cuándo la pinto para mi familia. "She asks me when I will paint my grandfather's house for my family."
39	Ella me pregunta/dice que esa maleta (con el vestido) (que) dónde la transporto por dos semanas. "She asks me where I will transport the suitcase with the dress for two weeks."
40	Ella me pregunta/dice que esa pintura (en el dormitorio) (que) cuándo la presto al museo moderno. "She asks me when I will loan the painting in the bedroom to the modern museum."
41	Ella me pregunta/dice que esa receta (de la ceremonia) (que) cuándo la cocino para mi hermano. "She asks me when I will cook the recipe from the ceremony for my brother."
42	Ella me pregunta/dice que esa semilla (en la cocina) (que) cuándo la siembro en el jardín. "She asks me when I will plant the kitchen seed in the garden."
43	Ella me pregunta/dice que esa vitamina (para los músculos) (que) cuándo la consumo durante el día. "She asks me when during the day I will consume the vitamin for my muscles."
44	Ella me pregunta/dice que ese aeropuerto (en la provincia) (que) cuándo lo describo para el conductor. "She asks me when I will describe the provincial airport to the driver."
45	Ella me pregunta/dice que ese arquitecto (de la alcaldía) (que) cuándo lo ayudo con el edificio. "She asks me when I will help the architect of the mayor's office with the building."
46	Ella me pregunta/dice que ese ensayo (sobre mi historia) (que) cuándo lo publico en una revista.

	"She asks when I will publish the essay about my history in the magazine."
	Ella me pregunta/dice que ese regalo (de la herencia) (que) dónde lo escondo en la casa.
47	"She asks me where in the house I will hide the inheritance gift."
	Ella me pregunta/dice que ese resultado (de la competencia) (que) dónde lo celebro con la familia.
48	"She asks me where I will celebrate the competition result with my family."

References

- Aarts, Alexander. 2015. Estimating the reproducibility of psychological science. *Science* 349: 943-950.
- Abutalebi, Jubin, and Pasquale Anthony Della Rosa. 2008. Imaging technologies. In *The Blackwell guide to research methods in bilingualism and multilingualism*. Edited by Li Wei and Melissa Moyer. John Wiley & Sons, pp. 132-157.
- Alarcón, Irma. 2011. Spanish gender agreement under complete and incomplete acquisition: Early and late bilinguals' linguistic behavior within the noun phrase. *Bilingualism: Language and Cognition* 14: 332–50.
- Alequín, Héctor Aponte, and Luis Ortiz López. 2015. Variación dialectal e interfaz sintáctica/semántica/pragmática: La anteposición de sujetos en cláusulas subordinadas en infinitivo. *Spanish in Context* 12: 396-418.
- Alfaraz, Gabriela. 2017. The variation of subject pronouns over time in Cuban Spanish. In *Cuban Spanish Dialectology: Variation, Contact, and Change*. Edited by Alejandro Cuza. Georgetown University Press, pp. 81-96.
- Altmann, Gerry, and Yuki Kamide. 1999. Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition* 73: 247-264.
- Altmann, Gerry, and Yuki Kamide. 2007. The real-time mediation of visual attention by language and world knowledge: Linking anticipatory (and other) eye movements to linguistic processing. *Journal of memory and language* 57: 502-518.
- Arnold, Jennifer, Janet Eisenband, Sarah Brown-Schmidt, and John Trueswell. 2000. The rapid use of gender information: Evidence of the time course of pronoun resolution from eyetracking. *Cognition* 76: B13–26.
- Audacity Team. 1999-2019. Audacity®: Free Audio Editor and Recorder. Version 2.0.3 retrieved 2014 from <https://audacityteam.org/>.
- Baayen, Harald, Douglas Davidson, and Douglas Bates. 2008. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of memory and language* 59: 390-412.

- Baayen, Harald, and Petar Milin. 2010. Analyzing reaction times. *International Journal of Psychological Research* 3: 12-28.
- Baddeley, Alan. 2000. The episodic buffer: a new component of working memory?. *Trends in cognitive sciences* 4: 417-423.
- Baddeley, Alan. 2003. Working memory: looking back and looking forward. *Nature reviews neuroscience* 4: 829-839.
- Baddeley, Alan. 2006. Working memory: An overview. *Working memory and education*, 1-31. Academic Press.
- Baddeley, Alan. 2007. Working memory, thought, and action. Oxford University Press.
- Baddeley, Alan, and Hitch, G. (1974). Working memory. *Psychology of learning and motivation* 8: 47-89. Academic press.
- Bailey, Karl, and Fernanda Ferreira. 2003. Disfluencies affect the parsing of garden-path sentences. *Journal of Memory and Language* 49: 183–200.
- Barton, Stephen, and Anthony Sanford. 1993. A case study of anomaly detection: Shallow semantic processing and cohesion establishment. *Memory & Cognition* 21: 477-487.
- Bosque, Ignacio, and Javier Gutiérrez-Rexach. 2009. *Fundamentos de sintaxis formal*. Ediciones Akal.
- Bosque, Ignacio, Violeta Demonte, Lázaro Carreter, María Victoria, Pavón Lucero, and Real Academia Española. 1999. *Gramática descriptiva de la lengua española*. Madrid: Espasa Calpe.
- Bowles, Melissa. 2011. Exploring the role of modality: L2-heritage learner interactions in the Spanish language classroom. *Heritage Language Journal* 8: 30–65.
- Brovetto, Claudia. 2002. Spanish clauses without complementizer. In *Current issues in Romance languages: Selected Proceedings of the 29th Linguistic Symposium on Romance Languages*. Edited by Teresa Satterfield, Christina Tortora, and Diana Cresti. Amsterdam & Philadelphia: John Benjamins, pp. 33-46
- Bruhn de Garavito, Joyce. 2002. Verb raising in Spanish: A comparison of early and late bilinguals. In *Proceedings of the 26th annual Boston University Conference on*

- Language Development*. Edited by Barbora Skarabela, Sarah Fish, and Anna Do. Somerville, MA: Cascadilla Press, pp. 84–94.
- Camacho, José. 2013. *Null subjects* (No. 137). Cambridge University Press.
- Casasanto, Laura, and Ivan Sag. 2008. The Advantage of the Ungrammatical. In *Proceedings of the 30th Annual Conference of the Cognitive Science Society*. Edited by Bradley Love, Ken McRae, and Vladimir Sloutsky. Austin, TX: Cognitive Science Society, pp. 601-606.
- Christensen, Rune Haubo. 2015. Analysis of ordinal data with cumulative link models—estimation with the R-package ordinal. R-version 3.6.3, pp. 1-31.
- Christianson, Kiel, Andrew Hollingworth, John Halliwell, and Fernanda Ferreira. 2001. Thematic roles assigned along the garden path linger. *Cognitive psychology* 42: 368-407.
- Collins, Harry. 2016. Reproducibility of experiments: experimenters’ regress, statistical uncertainty principle, and the replication imperative. In *Reproducibility: Principles, Problems, and Prospects*. Edited by Harald Atmanspacher and Sabine Maasen. New York: Wiley, pp. 65–82.
- Cuza, Alejandro. 2013. Crosslinguistic influence at the syntax proper: Interrogative subject–verb inversion in heritage Spanish. *International Journal of Bilingualism* 17: 71–96.
- Cuza, Alejandro. 2016. The status of interrogative subject–verb inversion in Spanish-English bilingual children. *Lingua* 180: 124-138.
- Cuza, Alejandro. 2017. *Cuban Spanish Dialectology: Variation, Contact and Change*. Georgetown University Press.
- Cuza, Alejandro, and Joshua Frank. 2011. Transfer effects at the syntax-semantics interface: The case of double-*que* questions in heritage Spanish. *The Heritage Language Journal* 8: 66-89.
- Cuza, Alejandro, and Joshua Frank. 2015. On the Role of Experience and Age-Related Effects: Evidence from the Spanish CP. *Second Language Research* 31: 3-28.

- Cuza, Alejandro, Ana Teresa Pérez-Leroux, and Liliana Sánchez. 2013. The role of semantic transfer in clitic drop among simultaneous and sequential Chinese-Spanish bilinguals. *Studies in Second Language Acquisition* 35: 93-125.
- Cuza, Alejandro, and Rocío Pérez-Tattam. 2016. Grammatical gender selection and phrasal word order in child heritage Spanish: A feature re-assembly approach. *Bilingualism: Language and Cognition* 19: 50-68.
- Dabrowska, Ewa. 2012. Different speakers, different grammars: Individual differences in native language attainment. *Linguistic Approaches to Bilingualism* 2: 219-253.
- Daneman, Meredyth, and Patricia Carpenter. 1980. Individual differences in working memory and reading. *Journal of Memory and Language* 19: 450-466.
- Davies, Mark. 2017. The new 4.3 billion word NOW corpus, with 4-5 million words of data added every day. In *The 9th International Corpus Linguistics Conference*.
- de Prada Pérez, Ana. 2009. Subject expression in Minorcan Spanish: Consequences of contact with Catalan. Ph.D. dissertation, Pennsylvania State University, State College, PA, USA.
- Demonte, Violeta, and Olga Fernández-Soriano. 2009. Force and finiteness in the Spanish complementizer system. *Probus* 21: 23-49.
- Demonte, Violeta, and Olga Fernández-Soriano. 2014. Evidentiality and illocutionary force: Spanish matrix *que* at the syntax-pragmatics interface. In *Left sentence peripheries in Spanish: Diachronic, variationist, and typological perspectives*. Edited by Andreas Dufter. Amsterdam: John Benjamins, pp. 217-252.
- Doya, Kenji, Shin Ishii, Alexandre Pouget, and Rajesh Rao. 2007. *Bayesian brain: Probabilistic approaches to neural coding*. MIT press.
- E-Prime 2.0 software. Psychology Software Tools, Pittsburgh, PA.
- Echeverría, Carlos, and Javier López Seoane. 2019. Re complementation as constrained CP resetting. Paper presented at the Association of Linguistics and Philology of Latin America (Alfalito), Queens College, City University of New York, New York City, NY, USA, September 20.

- Edelman, Shimon, and Morten Christiansen. 2003. How seriously should we take minimalist syntax? A comment on Lasnik. *Trends in Cognitive Science* 7: 60-61.
- Erker, Daniel., Eduardo Ho-Fernández, Ricardo Otheguy, and Naomi Lapidus Shin. 2017. Continuity and change in Spanish among Cubans in New York: A study of subject placement with finite verbs. *Cuban Spanish Dialectology: Variation, Contact, and Change*. Edited by Alejandro Cuza. Georgetown University Press, pp. 63-80.
- Etxepare, Ricardo. 2010. From hearsay evidentiality to samesaying relations. *Lingua* 120: 604-627.
- Ferreira, Fernanda. 2003. The misinterpretation of noncanonical sentences. *Cognitive Psychology* 47: 164–203.
- Ferreira, Fernanda. 2005. Psycholinguistics, formal grammars, and cognitive science. *Linguistic Review* 22: 365-380.
- Ferreira, Fernanda, Vittoria Ferraro, and Karl Bailey. 2002. Good-enough representations in language comprehension. *Current Directions in Psychological Science* 11: 11–15.
- Ferreira, Fernanda, and John Henderson. 1991. Recovery from misanalyses of garden-path sentences. *Journal of Memory and Language* 30: 725-745.
- Ferreira, Fernanda, and John Henderson. 1990. Use of verb information in syntactic parsing: Evidence from eye movements and word-by-word self-paced reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 16: 555-568.
- Ferreira, Fernanda, Ellen Lau, and Karl Bailey. 2004. Disfluencies, language comprehension, and tree adjoining grammars. *Cognitive Science* 28: 721-749.
- Ferreira, Fernanda, and Nikole Patson. 2007. The ‘good enough’ approach to language comprehension. *Language and Linguistics Compass* 1: 71-83.
- Fontana, Josep. 1993. Phrase structure and the syntax of clitics in the history of Spanish. Ph.D. dissertation, University of Pennsylvania, Philadelphia, PA, USA.
- Frank, Joshua. 2013. Derivational complexity effects in bilingual adults: Instances of interrogative inversion in Spanish. In *Selected Proceedings of the 16th Hispanic*

- Linguistics Symposium*. Edited by Jennifer Cabrelli, Gillian Lord, Ana de Prada Pérez, and Jessi Aaron. Somerville: Cascadilla Proceedings Project, pp. 143-55.
- Frank, Joshua. 2016. On the grammaticality of recomplementation in Spanish. In *Inquiries in Hispanic Linguistics: From theory to empirical evidence*. Edited by Alejandro Cuza, Lori Czerwionka, and Daniel Olson. Amsterdam: John Benjamins, pp. 39-52.
- Frank, Joshua. Submitted. The effect of silence and distance: Evidence from recomplementation in US heritage Spanish.
- Frank, Joshua, and Almeida Jacqueline Toribio. 2017. Recomplementation as an unexplored locus of dialectal variation: The status of reduplicative que in Cuban Spanish. *Cuban Spanish Dialectology: Variation, Contact, and Change*. Edited by Alejandro Cuza. Georgetown University Press, pp. 119-133.
- Fraser, Bruce. 1999. What are discourse markers? *Journal of Pragmatics* 31. 931–952.
- Frazier, Lyn. 1987. Sentence processing: A tutorial review. *Attention and performance XII*. Edited by Max Coltheart. Hillsdale, NJ: Lawrence Erlbaum, pp. 559–586.
- Frazier, Lyn., and Janet Fodor. 1978. The sausage machine: A new two-stage parsing model. *Cognition* 6: 291-325.
- Frazier, Lyn, and Keith Rayner. 1982. Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive psychology* 14: 178-210.
- Giancaspro, David. 2017. Heritage speakers' production and comprehension of lexically- and contextually-selected subjunctive mood morphology. Ph.D. dissertation, Rutgers University, New Brunswick, NJ, USA.
- Gibson, Edward. 1998. Linguistic complexity: Locality of syntactic dependencies. *Cognition* 68: 1–76.
- Gibson, Edward. 2000. The dependency locality theory: A distance-based theory of linguistic complexity. In *Image, Language, Brain*. Edited by Alec Marantz, Yasushi Miyashita, and Wayne O'Neil. MIT Press, pp. 95-126.

- Gibson, Edward, and Evelina Fedorenko. (2010a). Weak quantitative standards in linguistics research. *Trends in Cognitive Sciences* 14: 233-234.
- Gibson, Edward, and Evelina Fedorenko. (2010b). The need for quantitative methods in syntax and semantics research. *Language and Cognitive Processes* 28: 88-124.
- González i Planas, Francesc. 2014. On quotative recomplementation: Between pragmatics and morphosyntax. *Lingua* 146: 39-74.
- Grodner, Daniel, and Edward Gibson. 2005. Consequences of the serial nature of linguistic input for sentential complexity. *Cognitive Science* 29: 261–291.
- Gupton, Timothy. 2010. The Syntax Information Structure Interface: Subjects and Clausal Word Order in Galician. Ph.D. dissertation, University of Iowa, Iowa City, IA, USA.
- Hale, John. 2001. A probabilistic early parser as a psycholinguistic model. In *Proceedings of the Second Meeting of the North American Chapter of the Association for Computational Linguistics*. Association for Computational Linguistics, pp. 159–66.
- Hamayan, Else, Barbara Markman, Susanne Pelletier, and Richard Tucker. 1978. Differences in performance in elicited imitation between French monolingual and English-speaking bilingual children. *International Review of Applied Linguistics in Language Teaching* 16: 330-339.
- Hansen, Maj-Britt Mosegaard. 1998. The semantic status of discourse markers. *Lingua* 104. 235–260.
- Huetig, Falk, and Esther Janse. 2016. Individual differences in working memory and processing speed predict anticipatory spoken language processing in the visual world. *Language, Cognition and Neuroscience* 31: 80-93.
- Hulstijn, Jan. 2005. Theoretical and empirical issues in the study of implicit and explicit second-language learning: Introduction. *Studies in second language acquisition* 27: 129-140.
- Iatridou, Sabine, and Anthony Kroch. 1992. The Licensing of CP-recursion and its relevance to the Germanic verb-second phenomenon. *Working Papers in Scandinavian Linguistics* 50: 1-25.

- Isabelli, Casilde. 2004. The acquisition of the null subject parameter properties in SLA: Some effects of positive evidence in a naturalistic learning context. *Hispania* 87: 150–162.
- Janse, Esther, and Alexandra Jesse. 2014. Working memory affects older adults' use of context in spoken-word recognition. *The Quarterly Journal of Experimental Psychology* 67: 1842-1862.
- Jegerski, Jill. 2014. Self-paced reading. *Research methods in second language psycholinguistics*. Edited by Jill Jegerski and Bill VanPatten. New York: Routledge, pp. 20-49.
- Jurafsky, Dan. 1996. A probabilistic model of lexical and syntactic access and disambiguation. *Cognitive science* 20: 137-194.
- Just, Marcel., and Patricia Carpenter. 1992. A capacity theory of comprehension: individual differences in working memory. *Psychological review* 99: 122-149.
- Kamide, Yuki, Gerry Altmann, and Sarah Haywood. 2003. The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and language* 49: 133-156.
- Keating, Gregory, Jill Jegerski, and Bill VanPatten. 2016. Online processing of subject pronouns in monolingual and heritage bilingual speakers of Mexican Spanish. *Bilingualism: Language and Cognition* 19: 36–49.
- Kempchinsky, Paula. 2013. CLLD as a Window into the Left Periphery. *Selected Proceedings of the 15th Hispanic Linguistics Symposium (HLS 2011)*, ed. by C. Howe et al. Somerville, Mass.: Cascadilla Proceedings Project.
- Kim, Ji-Hye, Silvina Montrul, and James Yoon. 2010. Dominant language influence in acquisition and attrition of binding: Interpretation of the Korean reflexive caki. *Bilingualism: Language and Cognition* 13: 73–84.
- Kuperberg, Gina, and T. Florian Jaeger. 2016. What do we mean by prediction in language comprehension?. *Language, cognition and neuroscience* 31: 32-59.
- Lastra, Yolanda, and Pedro Butragueño. 2015. Subject pronoun expression in oral Mexican Spanish. *Subject pronoun expression in Spanish: A cross-dialectal perspective*.

- Edited by Ana Carvalho, Rafael Orozco, and Naomi Lapidus Shin. Georgetown University Press, pp. 39-57.
- Ledgeway, Adam. 2000. *A comparative syntax of the dialects of southern Italy. A minimalist approach*. Boston: Blackwell.
- Lewis, Richard. 2000. Falsifying serial and parallel parsing models: Empirical conundrums and an overlooked paradigm. *Journal of Psycholinguistic Research* 29: 241–248.
- Levy, Roger. 2008. Expectation-based syntactic comprehension. *Cognition* 106: 1126–77.
- Lipski, John. 1977. Preposed subjects in questions: Some considerations. *Hispania* 60: 61-67.
- López, Luis. 2009. Ranking the linear correspondence axiom. *Linguistic Inquiry* 40: 239-276.
- Ludlow, Peter. 2011. *The philosophy of generative linguistics*. Oxford University Press.
- MacDonald, M. C. (2013). Distributional information in language comprehension, production, and acquisition: Three puzzles and a moral. In *The emergence of language* (pp. 195-214). Psychology Press.
- MacDonald, Maryellen, Marcel Just, and Patricia Carpenter. 1992. Working memory constraints on the processing of syntactic ambiguity. *Cognitive Psychology* 24: 56–98.
- Martín-González, Javier. 2002. The Syntax of sentential negation in Spanish. Ph.D. dissertation, Harvard University, Cambridge, MA, USA.
- Martínez-Sanz, Cristina. 2011. Null and overt subjects in a variable system: The case of Dominican Spanish. Ph. D dissertation, University of Ottawa, Canada.
- Marull, Crystal. 2017a. Processing (In) Efficiency in the Second Language: Linguistic Experience and Cognitive Effects on Morphosyntactic Predictions. Ph.D. dissertation, Rutgers University, New Brunswick, NJ, USA.
- Marull, Crystal. 2017b. Second language processing efficiency: Experience and cognitive effects on L2 morphosyntactic integration and anticipation. In *Proceedings of the 41st Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press, pp. 466-480.

- Mascarenhas, Salvador. 2007. *Complementizer Doubling in European Portuguese*. Amsterdam/New York: ILLC/NYU.
- Mitchell, Don, Fernando Cuetos, Martin Corley, and Marc Brysbaert. 1995. Exposure-based models of human parsing: Evidence for the use of coarse-grained (nonlexical) statistical records. *Journal of Psycholinguistic Research* 24: 469-488.
- Montrul, Silvina. 2002. Incomplete acquisition and attrition of Spanish tense/aspect distinctions in adult bilinguals. *Bilingualism: Language and Cognition* 5: 39-68.
- Montrul, Silvina. 2004. Subject and object expression in Spanish heritage speakers: A case of morphosyntactic convergence. *Bilingualism: Language and Cognition* 7: 125-42.
- Montrul, Silvina. 2008. *Incomplete acquisition in bilingualism: Re-examining the age factor*. Amsterdam: John Benjamins.
- Montrul, Silvina. 2010. How similar are L2 learners and heritage speakers? Spanish clitics and word order. *Applied Psycholinguistics* 31: 167-207.
- Montrul, Silvina. 2011. Morphological errors in Spanish second language learners and heritage speakers. *Studies in Second Language Acquisition* 33: 155-161.
- Montrul, Silvina. 2016. *The acquisition of heritage languages*. Cambridge: Cambridge University Press.
- Montrul, Silvina, Rebecca Foote, and Silvia Perpiñán. 2008. Knowledge of Wh-movement in Spanish L2 learners and heritage speakers. In *Selected Proceedings of the 10th Hispanics Linguistics Symposium*. Edited by Joyce Bruhn de Garavito and Elena Valenzuela. Somerville: Cascadia Proceedings Project, pp. 93-106.
- Montrul, Silvina, and Roumyana Slabakova. 2003. Competence similarities between native and near-native speakers: An investigation of the Preterite/Imperfect contrast in Spanish. *Studies in Second Language Acquisition* 25: 351-98.
- Müller, Natascha, and Aafke Hulk. 2001. Crosslinguistic influence in bilingual language acquisition: Italian and French as recipient languages. *Bilingualism: Language and Cognition* 4: 1-21.

- O'Grady, William. 2010. An emergentist approach to syntax. *The Oxford handbook of linguistic analysis*, pp. 257-83.
- Orozco, Rafael. 2015. Pronominal variation in Colombian costeño Spanish. *Subject pronoun expression in Spanish: A cross-dialectal perspective*, pp. 17-37.
- Ortega-Santos, I. (2020a). Dialect distance and data assessment in Chilean Venezuelan & Puerto Rican Spanish. ResearchGate.net.
- Ortega-Santos, I. (2020b). Reliability in theoretical syntax: recent vs. classic high impact data and language choice. ResearchGate.net.
- Ortiz-López, Luis. 2009. El español del Caribe: orden de palabras a la luz de la interfaz léxico-sintáctica y sintáctico-pragmática. *Revista internacional de lingüística iberoamericana* 14: 75-93.
- Ortiz-López, Luis, Ashlee Dauphinais, and Héctor Aponte Alequín. 2017. Cuban Spanish: Is it a null subject parameter dialect?. *Cuban Spanish Dialectology: Variation, Contact, and Change*. Edited by Alejandro Cuza. Georgetown University Press, pp. 97-117.
- Otheguy, Ricardo, and Ana Celia Zentella. 2012. *Spanish in New York: Language Contact, Dialect Leveling, and Structural Continuity*. New York: Oxford University Press.
- Otheguy, Ricardo, Ana Celia Zentella, and David Livert. 2007. Language and dialect contact in Spanish in New York: Toward the formation of a speech community. *Language*, 770-802.
- Paoli, Sandra. 2006. The fine structure of the left periphery: COMPs and subjects: evidence from Romance. *Lingua* 117: 1057-79.
- Pascual y Cabo, Diego, and Jason Rothman. 2012. The (il) logical problem of heritage speaker bilingualism and incomplete acquisition. *Applied linguistics* 33: 450-55.
- Perez-Cortes, Silvia. 2016. Acquiring obligatory and variable mood selection: Spanish heritage speakers and L2 learners' performance in desideratives and reported speech contexts. Ph.D. dissertation, Rutgers University, New Brunswick, NJ, USA.

- Pérez-Leroux, Ana Teresa, Alejandro Cuza, and Danielle Thomas. 2011. Clitic Placement in Spanish–English Bilingual Children. *Bilingualism: Language and Cognition* 14: 221-232.
- Plann, Susan. 1982. Indirect Question in Spanish. *Linguistics Inquiry* 13: 297-312.
- Polinsky, Maria. 2011. Reanalysis in adult heritage language: A case for attrition. *Studies in Second Language Acquisition* 33: 305–28.
- Polinsky, Maria. 2018. *Heritage languages and their speakers*. Cambridge: Cambridge University Press.
- Polinsky, Maria, and Gregory Scontras. 2020. Understanding heritage languages. *Bilingualism: Language and Cognition* 23: 4-20.
- Potowski, Kim, Jill Jegerski, and Kara Morgan-Short. 2009. The effects of instruction on linguistic development in Spanish heritage language speakers. *Language Learning* 59: 537–79.
- Peirce, Jonathan, Jeremy Gray, Sol Simpson, Michael MacAskill, Richard Höchenberger, Hiroyuki Sogo, Erik Kastman, and Jonas Lindeløv. 2019. PsychoPy2: Experiments in behavior made easy. *Behavior research methods* 51: 195-203.
- R Core Team. 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Radford, Andrew. 2013. The complementiser system in spoken English. In *Information Structure and Agreement*. Edited by Victoria Camacho-Taboada, Ángel Jiménez-Fernández, Javier Martín-González and Mariano Reyes-Tejedor. Amsterdam: John Benjamins, pp. 11-54.
- Radford, Andrew. 2018. *Colloquial English: Structure and variation*. Cambridge: Cambridge University Press.
- Radloff, Carla, and Daniel Hallberg. 1991. Sentence Repetition Testing for Studies of Community Bilingualism. Presented at the Summer Institute of Linguistics, Dallas, Texas, and Arlington, Texas.

- Reynolds, Douglas. 2009. Gaussian Mixture Models. *Encyclopedia of biometrics*. Edited by Stan Li and Anil Jain. USA: Springer, pp. 659-663.
- Riccelli, Adrian. 2018. Espero estén todos: The distribution of the null subordinating complementizer in two varieties of Spanish. *Language Variation and Contact-Induced Change: Spanish across space and time*. Edited by Jeremy King and Sandro Sessarego. John Benjamins, pp.299-333.
- Rivero, M. 1980. On Left-Dislocation and Topicalization in Spanish. *Linguistic Inquiry* 11: 363-395.
- Rivero, M. 1994. On Indirect Questions, Commands, and Spanish Quotative Que. *Linguistic Inquiry* 25: 547-554.
- Rizzi, Luigi. 1997. The fine structure of the left periphery. In *Elements of Grammar*. Edited by Liliane Haegeman. Dordrecht: Kluwer Academic Publishers, pp.281-337.
- Rizzi, Luigi. 2013. Notes on cartography and further explanation. *Probus* 25: 197 – 226.
- Rodríguez-Ramalle, Teresa. 2003. *La gramática de los adverbios en -mente o cómo expresar maneras, opiniones y actitudes a través de la lengua*. Madrid: Ediciones de la Universidad Autónoma de Madrid.
- Roland, Douglas, Frederic Dick, and Jeffrey Elman. 2007. Frequency of English grammatical structures: A corpus analysis. *Journal of Memory and Language* 57: 348–379.
- Rönnerberg, Jerker, Thomas Lunner, Adriana Zekveld, Patrik Sörqvist, Henrik Danielsson, Björn Lyxell, Örjan Dahlström, Carine Signoret, Stefan Stenfelt, M. Kathleen Pichora-Fuller, and Mary Rudner. 2013. The Ease of Language Understanding (ELU) model: theoretical, empirical, and clinical advances. *Frontiers in systems neuroscience*, 7(31).
- Sagarra, Nuria., and Julia Herschensohn. 2013. Processing of gender and number agreement in late Spanish bilinguals. *International Journal of Bilingualism* 17: 607-627.
- Sánchez, Liliana. 2019. Bilingual Alignments. *Languages* 4, 82. <https://doi.org/10.3390/languages4030069>

- Sanford, Anthony, and Patrick Sturt. 2002. Depth of processing in language comprehension: Not noticing the evidence. *Trends in cognitive sciences* 6: 382-386.
- Scontras, Gregory, Maria Polinsky, and Zuzanna Fuchs. 2018. In support of representational economy: Agreement in heritage Spanish. *Glossa: A journal of general linguistics*, 3.
- Scontras Gregory, Maria Polinsky, C. Y. Edwin Tsai, and Kenneth Mai. 2017. Cross-linguistic scope ambiguity: When two systems meet. *Glossa: A journal of general linguistics*, 2.
- Silva-Corvalán, Carmen. 1993. On the permeability of grammars: Evidence from Spanish and English contact. In *Linguistic Perspectives on Romance Languages*. Edited by William Ashby, Marianne Mithun, and Giorgio Perissinotto. Amsterdam: John Benjamins, pp. 19–43.
- Silva-Corvalán, Carmen. 1994. *Language contact and change. Spanish in Los Angeles*. Oxford: Oxford University Press.
- Sorace, Antonella. 2000. Differential effects of attrition in the L1 syntax of near-native L2 speakers. In *Proceedings of the 24th Boston University Conference on Language Development*. Edited by Catherine Howell, Sarah Fish, and Thea Keith-Lucas. Somerville, MA: Cascadilla Press, pp. 719–25.
- Sorace, Antonella. 2004. Native language attrition and developmental instability at the syntax-discourse interface: Data, interpretations and methods. *Bilingualism: Language and Cognition* 7: 143–5.
- Sorace, Antonella. 2005. Selective optionality in language development. In *Syntax and variation: Reconciling the biological and the social*. Edited by Leonie Cornips and Karen Corrigan. Amsterdam: John Benjamins, pp. 46–111.
- Sorace, Antonella. 2011. Pinning down the concept of “interface” in bilingualism. *Linguistic Approaches to Bilingualism* 1: 1-33.
- Sprouse, Jon, Carson Schütze, and Diogo Almeida. 2013. Assessing the reliability of journal data in syntax: Linguistic Inquiry 2001-2010. *Lingua* 134: 219-248.

- Staub, Adrian. 2010. Eye movements and processing difficulty in object relative clauses. *Cognition* 116: 71-86.
- Suñer, Margarita. 1992. Indirect questions and structure of CP: Some Consequences. *Current Studies in Spanish Linguistics*. Edited by Héctor Campos and Fernando Martínez-Gil. Washington DC: Georgetown University Press, pp. 283-312.
- Suñer, Margarita. 1993. About indirect questions and semi-questions. *Linguistics and Philosophy*, 45-77.
- Tabori, Andrea, Emily Mech, and Natsuki Atagi. 2018. Exploiting language variation to better understand the cognitive consequences of bilingualism. *Frontiers in psychology*, 9(1686).
- Toribio, Almeida Jacqueline. 2000. Setting parametric limits on dialectal variation in Spanish. *Lingua* 110: 315-341.
- Traxler, Matthew, Martin Pickering, and Charles Clifton Jr. 1998. Adjunct attachment is not a form of lexical ambiguity resolution. *Journal of Memory and Language* 39: 558-592.
- Tucker, Matthew, Ali Idrissi, and Diogo Almeida. 2015. Representing number in the real-time processing of agreement: self-paced reading evidence from Arabic. *Frontiers in psychology*, 6(347).
- Unsworth, Sharon. 2016. Quantity and quality of language input in bilingual language development. In *Lifespan perspectives on bilingualism*. Edited by Elena Nicoladis and Simona Montanari. Berlin: Mouton de Gruyter, pp. 136–196.
- Uriagereka, Juan. 1995. An F position in Western Romance. *Discourse Configurational Languages*. Edited by Katalin Kiss. Oxford: Oxford University Press, pp. 153-175.
- Van Gompel, Roger, Martin Pickering, and Matthew Traxler. 2001. Reanalysis in sentence processing: Evidence against current constraint-based and two-stage models. *Journal of Memory and Language* 45: 225-258.
- Villa-García, Julio. 2012. Re complementation and locality of movement in Spanish. *Probus* 24: 257-314.

- Villa-García, Julio. 2015. *The Syntax of Multiple-que Sentences in Spanish: Along the Left Periphery*. Amsterdam: John Benjamins.
- Villa-García, Julio. 2019. Re complementation in English and Spanish: Delineating the CP space. *Glossa: A journal of general linguistics*, 4.
- Warren, Tessa, and Edward Gibson. 2002. The influence of referential processing on sentence complexity. *Cognition* 85: 79-112.
- Yip, Virginia, and Stephen Matthews. 2009. Conditions on cross-linguistic influence in bilingual acquisition: The case of wh-interrogatives. Paper presented at the 7th International Symposium on Bilingualism, Utrecht, Netherlands, July.
- Zapata, Gabriela, Liliana Sánchez, and Almeida Jacqueline Toribio. 2005. Contact and contracting Spanish. *International Journal of Bilingualism* 9: 377-95.