University of Southampton
Faculty of Arts and Humanities
Department of Modern Languages and Linguistics

An Investigation into the Acquisition and Processing of the Subjunctive by English-speaking Second Language Learners of French

by

Amber Dudley

ORCID iD 0000-0003-2904-9150

Thesis submitted for the degree of Doctor of Philosophy (PhD) in Modern Languages

July 2020

University of Southampton

Abstract

Faculty of Arts and Humanities

Modern Languages and Linguistics

Thesis for the degree of Doctor of Philosophy (PhD) in Modern Languages

An Investigation into the Acquisition and Processing of the Subjunctive by English-speaking Second Language Learners of French

by Amber Dudley

Learning how to express grammatical mood (indicative/subjunctive) in a second language has been shown to be problematic for learners whose first language (e.g., English) does not fully grammaticalise this distinction. Based on these findings, the current thesis presents a series of eye-tracking studies investigating the extent to which second language learners can acquire target-like knowledge of the French subjunctive and apply it during real-time processing, particularly in contexts where the first and second language exhibit crosslinguistic differences. A secondary aim of this thesis is to explore whether this ability is influenced by proficiency, residence in a French-speaking country and reading goals.

A combination of eye movement and judgment data revealed that second language learners consistently exhibited target-like knowledge of the French subjunctive in obligatory contexts, but this knowledge did not always translate into target-like processing patterns. In particular, we found that the ability to process mood-modality mismatches depended, in large part, on the first language properties, proficiency and residence in a French-speaking country among the second language learners. In contrast, both first and second language speakers exhibited an absence of sensitivity to the discourse-pragmatic constraints of the polarity subjunctive. This ability, however, was not constrained by either proficiency and/or residence abroad among the second language group. Finally, reading goals played an important role in modulating second language sensitivity to verbal mood, with more target-like performance in tasks requiring a focus on form rather than function.

In a nutshell, the current thesis shows that target-like knowledge of the French subjunctive is attainable for second language learners, at least in obligatory contexts. However, multiple factors, including first language properties, residence abroad and residence goals, play a crucial role in modulating second language learners' ability to apply this knowledge during real-time processing. In contrast, it would appear that inconsistent input from the first language speakers prevents second language learners from acquiring target-like knowledge of the subjunctive in polarity contexts.

Table of Contents

Table of Tables

Table of Figures

Research Thesis: Declaration of Authorship

Print name: Amber Dudley

Title of thesis: An Investigation into the L2 Acquisition and Processing of the Subjunctive by English-Speaking Second Language Learners of French.

I declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signature: ....Date:.............................

Acknowledgements

Over the past three years, I have discovered that doing a PhD is most definitely not a solo endeavour. As much as I have enjoyed researching the subjunctive and learning about eye-tracking, there have been many individuals who have made the experience all the more enjoyable (and bearable!).

- Thank you first and foremost to my supervisor, Professor Roumyana Slabakova, for the exceptional support, expertise and guidance that you have generously offered me over the past six years, from my undergraduate dissertation to the final stages of this PhD. Thank you most of all for believing in my ability to be a linguist, a researcher and an academic. I am beyond thankful to have a supervisor like you.
- Thank you to all those who have given me feedback on earlier versions of this thesis. Particular thanks go to my second supervisor, Dr Jaine Beswick, and to my examiners at progression milestones, Professor Laura Domínguez and Dr Denis Drieghe. Your combined expertise has greatly improved the quality of this thesis.
- Thank you to my examiners, Dr Denis Drieghe and Professor Jason Rothman, for providing invaluable feedback during the viva and making it such an enjoyable experience.
- Thank you to James Turner for proofreading this thesis. I am eternally grateful.
- Thank you to the GenSLA research group at the University of Southampton. I have immensely enjoyed our time together over the years, both during reading groups and many memorable social occasions. Special thanks go to James Corbet, James Turner, Lewis Baker and Dr Eloi Puig-Mayenco for being the best sounding boards, for making the long hours spent in the office so entertaining and for always agreeing to pizza.
- Thank you to the Modern Languages and Linguistics department at the University of Southampton, my academic home for the past (nearly) nine years. I feel incredibly lucky to have been a member of this department. Special thanks go to Adrian Sewell for your infinite source of wisdom and for helping me to truly believe that I was as equally capable as my peers and to Professor James Minney for giving me the opportunity to perfect my answer to why the subjunctive is so fascinating thanks to your endless (and often quite insightful!) inquisitions.
- Thank you to the Second Language Acquisition and Bilingualism lab, in particular Professor Silvina Montrul and Professor Tania Ionin, for hosting me during my semester at the University of Illinois at Urbana-Champaign. Particular thanks go to Dr Sara Fernández Cuenca for helping me to get started with eye-tracking.
- Thank you to my friends and family for being there every step of the way, for helping me to remain (relatively!) sane and for providing welcome distractions in the form of videocalls, days out and weekends away. Special thanks go to Katie, Tasman, Hannah, Valentina, Isabel, Angharad and Maira. I couldn't wish for more supportive friends.
- Thank you to Dobby for being the most supportive and loving dog in the world.
- Thank you to every single individual who took part in this study. Without you, this study would not have been possible.

This PhD was generously funded by a 1+3 Scholarship from the UK Economic and Social Research Council (ESRC) Doctoral Training Centre at the University of Southampton (ES/J500161/1). The experimental work presented in this thesis was approved by the Faculty Ethics Committee (ERGO #46014 and #52953).

Definitions and Abbreviations

BIA+	Bilingual Interactive Activation Plus
Comp.	Comprehension
CP	Complement Phrase
DP	Determiner Phrase
FinP	Finite Phrase
ForceP	Force Phrase
FRH	Feature Reassembly Hypothesis
FTFA	Full Transfer/Full Access
GenSLA	Generative Second Language Acquisition
Ind.	Indicative
Judg.	Judgment
L1	First Language
L2	Second Language
M	Modal
MoodP	Mood Phrase
N	Number
Neg.	Negation
Ø	Null Element
PLAD	Parser as the Language Acquisition Device
POS	Poverty of the Stimulus
Prof.	Proficiency
QPT	Quick Placement Test
RAGE	Reduced Ability to Generate Expectations
SSH	Shallow Structure Hypothesis
Subj.	Subjunctive
T	Tense
TP	Tense Phrase
U	Uninterpretable
UG	Universal Grammar
V	Verb
VP	Verb Phrase

Chapter 1 Introduction

Formal approaches to second language (L2) development have typically used offline measures to examine the role of the first language (L1) on L2 comprehension. Much less is known, however, about how L1 grammatical knowledge, particularly L1/L2 differences in feature configurations, influences L2 online processing. Some scholars have proposed that L1/L2 differences play a deterministic role in modulating L2 learners' sensitivity to morphosyntax during real-time processing (e.g., the Full Transfer/Full Access/Full Parse Hypothesis, Dekydtspotter, Schwartz and Sprouse, 2006; the Morphological Congruency Hypothesis, Jiang *et al.*, 2011; the Unified Competition Model, MacWhinney, 2011; the Parser as the Language Acquisition Device, Dekydtspotter and Renaud, 2014), whereas other models do not attribute a significant role to L1 transfer, or at least remain agnostic on the subject (e.g., the Shallow Structure Hypothesis, Clahsen and Felser, 2006, 2018). It has even been argued that processing difficulties can hinder the language acquisition process itself (O'Grady, 2005; Sharwood Smith and Truscott, 2005; VanPatten, 2007).

The present thesis aims to test these proposals by examining the acquisition and processing of the subjunctive by English-speaking L2 learners of French in adulthood across a spectrum of proficiencies. The French subjunctive is a particularly interesting test case for assessing the development of sentence processing since it involves a long-distance dependency where mood choice in the rightmost clause requires the integration of syntactic, semantic and discourse information. This research thus builds on previous offline studies, which have shown that the subjunctive can be a source of persistent difficulty for L2 learners of French, even at the most advanced stages of proficiency (Howard, 2008; Bartning, Lundell and Hancock, 2012; Ayoun, 2013; McManus and Mitchell, 2015). No study to date, however, has used online methods, such as self-paced reading, eye-tracking or event-related potentials (ERPs), to investigate the L2 development of the French subjunctive. As we will discuss in the context of this thesis, eye-tracking has the potential to provide us with a more nuanced understanding of underlying linguistic knowledge.

1.1 Research Questions

The present thesis contributes to the existing body of research by examining the L2 acquisition and processing of the French subjunctive, a long-distance syntactic dependency, through the online paradigm of eye-tracking during reading. The use of eye-tracking provides a fined-grained measure of linguistic competence that does not necessarily induce task effects typically found in studies that rely on L2 learners' metalinguistic awareness. More specifically, this study aims to test the predictions of a processing-based extension to the Feature Reassembly Hypothesis (Lardiere, 2009). Such an extension would predict that the L1 morphosyntax modulates the processing of the subjunctive in the L2 at lower proficiency levels, but that target-like processing is ultimately attainable as learners become more proficient. Although the Feature Reassembly Hypothesis (Lardiere, 2009) has almost exclusively been a theory of L2 acquisition and not of processing, this study attempts to explore the parallel between the two. With this in mind, we advance the following research questions:

- Do English-speaking L2 learners of French exhibit offline and online sensitivity to the syntactic and semantic constraints of the subjunctive in obligatory contexts? In particular, to what extent is this sensitivity influenced by L1/L2 differences in feature configurations, proficiency and residence in a French-speaking country?

- Do English-speaking L2 learners of French exhibit offline and online sensitivity to the syntactic and discourse/pragmatic constraints of the subjunctive in contexts where the indicative and the subjunctive alternate? In particular, to what extent is this knowledge influenced by proficiency and residence in a French-speaking country?

- To what extent do secondary tasks (i.e., comprehension questions vs. judgment tasks) modulate online sensitivity to the syntactic, semantic and discourse/pragmatic constraints of the subjunctive?

In order to test these research questions, we conducted four studies. As we will discuss, the presence of the subjunctive in the English language has been widely debated in the literature. In light of this, Study 0 (Chapter 3) sought to establish the contexts in which native speakers of British English allow the subjunctive in their L1 grammar, using an acceptability judgment task.

Study 1 (Chapter 4) examined offline and online sensitivity to the subjunctive in obligatory contexts, using the results from Study 0 as a basis upon which to formulate concrete predictions relating to the influence of L1/L2 differences in feature configurations (i.e., the licensing contexts as determined by the lexical-semantic properties of the matrix predicate) on L2 knowledge and processing of the subjunctive. A secondary aim of this study was to further explore how proficiency (as conceptualised through lexical knowledge) and residence in a French-speaking country modulated this pattern. This study included the following tasks: a mood morphology production task and two eye-tracking during reading tasks. The first of the eye-tracking tasks instructed participants to read for comprehension, whereas the second required participants to assess the grammaticality of each sentence.

Study 2 (Chapter 5) investigated offline and online sensitivity to the subjunctive in contexts where the indicative and the subjunctive can alternate without ungrammaticality, but where such an alternation depends on discourse-pragmatic constraints. As with Study 1, a secondary aim of this study was to further explore how proficiency (as conceptualised through lexical knowledge) and residence in a French-speaking country modulated this pattern. Similar to Study 1, participants completed two eye-tracking during reading tasks: the first involving reading for comprehension and the second reading for judgment.

As we mentioned above, both Study 1 and Study 2 included two eye-tracking tasks: reading for comprehension and reading for judgment. Study 3 (Chapter 6) combined the results from these two studies in order to investigate the extent to which secondary tasks (i.e., comprehension questions vs. judgment tasks) influenced L2 learners' online sensitivity to verbal mood manipulations.

1.2 Structure of the Thesis

This thesis is divided into six chapters. Chapter 2 lays the theoretical foundations and discusses current theories of L2 acquisition and sentence processing, with a particular focus on the use of eye-tracking as a methodology for sentence processing research. Chapters 3, 4, 5, 6 present the individual studies, as described above. In particular, each chapter contextualises the respective study within previous scholarly work, describes the experimental design and the results of each study and then explores the implications of these findings. Finally, Chapter 7 outlines the main findings alongside the broader implications of this thesis.

Chapter 2 Theoretical Framework

The aim of this chapter is to present current theories of L2 acquisition and language processing in order to situate the thesis within the wider field of L2 development. First, Section 2.1.1 describes current theories of L2 acquisition couched predominately within a generative framework, while Section 2.1.2 examines the role of input in L2 acquisition. Section 2.2 then introduces current theories of language processing, looking in particular at lexical, syntactic, semantic and discourse processing; the focus of Section 2.2.1 and Section 2.2.2 is on L1 and L2 processing, respectively. Building on this, Section 2.2.3 explores how processing research can inform language acquisition debates, particularly in terms of how learners move from one state of (linguistic) knowledge to another. Finally, Section 2.2.4 discusses the use of eye-tracking as a methodology in sentence processing research and its applicability to and relevance for research on L2 development.

2.1 Second Language Acquisition

2.1.1 Theories of Second Language Development

The dynamic nature of variation in both performance and competence among the L2 population is a phenomenon that has dominated the generative second language acquisition (GenSLA) research paradigm for more than two decades (Rothman and Slabakova, 2018). GenSLA researchers have undertaken extensive experimental work in an attempt to account for the observable and well documented differences in ultimate attainment, namely in the domain of linguistic competence and performance, between L1 and L2 speakers that Universal Grammar (UG)[1] accessibility alone cannot explain. This research paradigm has led to a proliferation of generatively grounded theories that seek to elucidate the deterministic factors contributing to L1/L2 differences, with the underlying assumption that adult L2 learners have unabridged access to UG. The theoretical debate in GenSLA has historically been polarised by two competing positions: the representation deficit view and the full functional representation

[1] Universal Grammar refers to “an innate form of knowledge specific to language” (Yang, 2004, p. 451) that, in a sense, represents a genetic blueprint of how human languages function. The Minimalist program (Chomsky, 1995, 1998, 2000) posits that the human language faculty consists of a universal computation system and a lexicon which encompasses lexical items composed of syntactic, semantic and phonological feature bundles.

approach. In this section, we present a selective overview of these debates and theories in order to provide a theoretical foundation for future discussion.

At one end of the spectrum, proponents of the representation deficit have largely argued that L1/L2 differences in linguistic competence and performance stem from an impairment to the language faculty. Theories subsumed under this approach include the Representation Deficit Hypothesis (Hawkins and Franceschina, 2004), the Failed Functional Features Representation Hypothesis (Hawkins and Chan, 1997) and the Interpretability Hypothesis (Hawkins and Hattori, 2006; Tsimpli and Dimitrakopoulou, 2007). The most recent formulation, the Interpretability Hypothesis, stipulates that the interpretability of functional features[2] plays a deterministic role in successful L2 acquisition, such that interpretable features (i.e., those that convey semantic meaning, such as plural marking on English nouns) should not, in theory, be problematic for L2 learners, but that the acquisition of new uninterpretable features (i.e., those that do not convey any semantic meaning, such as gender agreement on French determiners and adjectives) may be a source of residual optionality for L2 learners. Numerous studies have investigated the predictions formulated by this hypothesis, providing both evidence for (Franceschina, 2001, 2005; Tsimpli and Dimitrakopoulou, 2007) and against it (Domínguez, 2007; Foucart and Frenck-Mestre, 2012; Hopp, 2013; Roberts, Sagarra and Herschensohn, 2013; Leal Méndez and Slabakova, 2014).

At the other end of the spectrum, the full functional representation approach, as formalised by the Full Transfer-Full Access (FTFA) hypothesis (Schwartz and Sprouse, 1994, 1996), argues against an impairment to the language faculty, assuming that both interpretable and uninterpretable functional features are fully acquirable even if not activated in the L2 learners' L1 grammar. The FTFA hypothesis instead adopts the view that the initial L2 (interlanguage) grammar consists of all the formal features directly transferred from the L1 grammar and that adult L2 learners, despite possible maturational constraints, still have unabridged access to UG. However, this does not mean that complete success is always guaranteed. L1/L2 differences requires parameter resetting, a process contingent on adequate exposure to positive L2 evidence (Slabakova, 2015).

The theoretical underpinnings of the FTFA hypothesis have paved the way for modern GenSLA theories, including the Interface Hypothesis and the Feature Reassembly Hypothesis. The Interface Hypothesis, assuming the central tenets of the FTFA model, proposes that the

---

[2] In this thesis, we use functional features to refer to the "formal properties of syntactic objects which determine how they behave with respect to syntactic constraints and operations" (Svenonius, 2019, p. 1).

acquisition of narrow syntax (that is, uninterpretable features internal to the computational system) should not present L2 learners with any significant challenges. Difficulties, however, are predicted to occur when L2 learners are faced with linguistic properties that lie at the interface between syntax and other cognitive domains (namely, discourse-pragmatics). Sorace (2003, 2004, 2005, 2011, 2012, 2014) and colleagues (Sorace and Filiaci, 2006; Tsimpli and Sorace, 2006; Belletti, Bennati and Sorace, 2007; Sorace and Serratrice, 2009; Filiaci, Sorace and Carreiras, 2014), who adopt a modular view to language, argue for a dichotomous relationship between internal and external interfaces with respect to the extent of fossilisation and non-target-like behaviour (White, 2009, 2011). In their view, external interfaces (i.e., the interface between syntax and other cognitive domains such as discourse and/or pragmatics) are, in theory, more challenging than internal interfaces (i.e., the interface between syntax and other sub-modules of language, such as semantics, phonetics/phonology and/or morphology).

Explanations for such an account have predominately been couched in terms of the processing demands that interface conditions place on the parser[3] due to insufficient computational resources. Although the Interface Hypothesis has provoked a lively epistemological debate surrounding the acquirability of interface conditions, empirical support remains mixed, with many studies providing clear evidence in support of the successful acquisition of external interfaces by advanced L2 learners (Iverson, Kempchinsky and Rothman, 2008; Rothman, 2009; Slabakova and Ivanov, 2011; Ivanov, 2012; Slabakova, Kempchinsky and Rothman, 2012).

Other modern GenSLA theories, however, have built more directly on the theoretical foundations laid by the FTFA hypothesis. The Feature Reassembly Hypothesis (FRH), for example, does not predict that L1/L2 differences can be explained in terms of an impairment to the language faculty and/or interface conditions. Instead, differences are argued to arise as the result of the challenges faced when mapping and reassembling new functional features onto L2-specific bundles and lexical items (Lardiere, 2009). The FRH notes that unlike L1 acquisition, L2 learners have access to "an entrenched system of morphosyntactic features already assembled into lexical items" (Hwang and Lardiere, 2013, p. 58) at the initial stages of L2 development. For successful acquisition of formal features to occur, L2 learners must undergo

[3] Here, the parser refers to the module of the mind that is responsible for the morphosyntactic analysis of the linguistic input and that sub-serves sentence processing (Fodor, 1983, 2000). The parser is thought to generate "UG-sanctioned representations that must be licensed by language-specific information at each step" (Dekydtspotter and Renaud, 2014, p. 133).

two distinct stages. The first involves comparing and contrasting L1 and L2 features in an attempt to map, either accurately or inaccurately, L1 forms onto L2 forms. In the event of L1/L2 form-to-meaning mismatches, L2 learners must reconfigure and reassemble existing L1 feature sets into new L2-specific feature bundles and lexical items. It is hypothesised that linguistic properties involving complex L1/L2 mismatches should generate a greater degree of difficulty. This hypothesis has found empirical support in recent studies by Guijarro-Fuentes (2012), Gil and Marsden (2013), Hwang and Lardiere (2013), Cho and Slabakova (2014) and Domínguez, Arche and Myles (2017), among others.

In summary, this section has explored the main theoretical debates in the field of generative L2 acquisition, including the Interpretability Hypothesis, the Full Transfer-Full Access Hypothesis, the Interface Hypothesis and the Feature Reassembly Hypothesis. These theories have undoubtedly raised numerous epistemological and ontological questions relating to L1/L2 differences in ultimate attainment. Yet, there still remains a distinct lack of consensus surrounding the source of these difficulties. The current study seeks to explore these questions further by investigating the extent to which differences in L1/L2 featural compositions and interface conditions influence the successful acquisition of the French subjunctive.

2.1.2 The Role of Input

There is a general impression in the wider L2 research community that generative researchers have largely overlooked the role that input plays in L2 development. Slabakova and colleagues (2014) sought to address these claims, which focused on the importance of input within the generative paradigm for setting parameter values. The authors argued that learners must be exposed to abundant, unambiguous and consistent input for parameter setting and successful acquisition to occur. However, in a rebuttal to this paper, de Bot (2015, p. 263) stated that generative researchers were not aware of the fact that "input is not abundant, unambiguous, and consistent, but messy, inconsistent, and ambiguous". According to Rankin and Unsworth (2016, p. 564) such an argument is "somewhat strange as inconsistency and messiness in the input is one element in the generative assumption of the poverty of the stimulus (POS)".[4] It has therefore been suggested that generative studies should systematically investigate the "quantitative distribution of input properties" and the "qualitative properties of the input available to learners" (Rankin and Unsworth, 2016, p. 563).

[4] The poverty of the stimulus refers to the argument that the input is insufficient in accounting for a learner's knowledge of a particular linguistic feature (Lasnik and Lidz, 2016).

As stated in the introduction, a secondary aim of this thesis is to investigate the extent to which an extended residence in a French-speaking country modulates L2 sensitivity to verbal mood. Although such an aim does not explicitly address Rankin and Unsworth's recommendations regarding a systematic investigation of the quantitative and qualitative properties of the input, it does explore, to a certain extent, the role of input in generative approaches to L2 acquisition and processing by considering how the nature of L2 exposure impacts on ultimate attainment. The current sub-section aims to provide a brief overview of previous studies that have examined the influence of residence abroad with respect to L2 grammatical development. Note that most of these studies have almost exclusively adopted a usage-based approach.

Although many have argued that learners must be exposed to an abundance of unambiguous and consistent input, previous research suggests that study abroad only offers limited benefits and, in most cases, classroom instruction has the more beneficial impact (Howard and Schwieter, 2018). For example, Collentine (2004) studied the development of verbal, adjectival and nominal morphosyntactic features, focusing in particular on tense, aspect, agreement and gender in L2 Spanish, among a group of classroom learners and one semester study abroad learners. Findings based on oral proficiency interviews revealed that classroom learners were in fact stronger than the study abroad learners.

Isabelli-García (2010) investigated the L2 development of gender agreement on adjectival marking among classroom learners and one semester study abroad learners of L2 Spanish, using a grammaticality judgment task. This study did not detect any differences between the two learner groups. A similar finding has also been uncovered for L2 French learners. Freed, Lazar and So (2003) studied native speaker evaluations of learners' written production, focusing on syntactic subordinators, before and after a semester abroad, and found no difference between the two groups.

Some evidence of limited benefits, however, has been observed among L2 German speakers. Arnett (2013) examined the L2 use of accusative and dative case prepositions and auxiliary verbs among learners before and after a semester abroad, using a picture description task. Although the analysis did not reveal a significant difference between the study abroad learners and their classroom counterparts, there were more tokens of ditransitive clauses among the study abroad group.

While the studies presented above largely suggest that classroom instruction has the more beneficial impact, further research suggests that the type of linguistic feature can influence the effect of study abroad. For example, over the past 15 years, Howard has conducted a series of studies looking at the influence of study abroad on the L2 acquisition of verbal morphology. In particular, Howard (2005a) found evidence of significant developmental gains among study abroad learners with respect to the perfective interpretation of the *passé composé* and the *imparfait*, but not with the habitual and progressive use of the *imparfait*. Further gains from study abroad were observed with future time forms of the periphrastic and inflected future and the conditional (Howard, 2012b). In contrast, lower frequency structures, such as the subjunctive and pluperfect for the expression of reverse order, proved to be more challenging for L2 learners, regardless of their study abroad experience (Howard, 2005b, 2012a).

A similar study by Grey, Cox, Serfaini and Sanz (2015) used a grammatical judgment task and reaction time data to examine development in L2 Spanish learners before and after a short five-week study abroad programme. Findings revealed that learners' judgments were more accurate for word order and number agreement, but not for gender agreement, and that reaction times became quicker over the course of the study.

Directly relevant to the current thesis is a study by McManus and Mitchell (2015), which investigated how subjunctive use develops over the course of a residence abroad experience. The authors found that although overall development was limited, learners, particularly those of a lower proficiency, exhibited more distinct gains in obligatory than variable contexts (e.g., with epistemic verbs in negated clauses). Howard (2008) presented similar findings using sociolinguistic interviews with English-speaking university learners (n = 18). Results showed that formal instruction improved subjunctive use, but that this development was restricted to a limited number of triggers and forms. Study abroad did, however, appear to widen the range of forms used among L2 learners, potentially due to increased exposure to naturalistic input.

To our knowledge, few studies have considered how study abroad might impact on L2 processing patterns. A study by LaBrozzi (2012) is a notable exception. The author used eye movement data to examine the use of morphological and lexical cues by study abroad and classroom learners of L2 Spanish when assigning temporal reference. He predicted that classroom learners would rely more on lexical than morphological cues given the lack of morphological richness in the L1 (English), whereas study abroad learners would depend on both cue types. These predictions were confirmed, suggesting that study abroad experience modulates processing patterns.

In summary, the current sub-section has shown that study abroad experience only offers limited benefits to L2 learners and, in most cases, classroom instruction has the more beneficial impact. It is important to recognise that some linguistic properties may be more sensitive to the effects of study abroad than others. However, future research is needed to further expand our understanding of how study abroad influences L2 development, particularly in terms of processing patterns.

2.2 Sentence Processing

Recent years have witnessed an emergence of generatively grounded L2 studies exploring processing-based accounts for L1/L2 differences in ultimate attainment. In particular, these studies have sought to examine whether non-native sentence processing is fundamentally (that is, quantitatively, or perhaps even qualitatively) different from native sentence processing, as well as the factors contributing to these differences, particularly in cases where learners may demonstrate native-like grammatical knowledge, but non-native-like processing profiles (Felser and Cunnings, 2012). For example, VanPatten and Jegerski (2010, p. 8) recognise that the acquisition of the relevant grammatical knowledge "does not automatically entail the processing of something related". Exploring L1/L2 processing differences can therefore help us to develop a more fine-grained understanding of L2 knowledge. The aim of this section is to discuss theories of L2 sentence processing in more depth. However, before pursuing this any further, we must first consider how language comprehension unfolds in our first language.

2.2.1 L1 Sentence Processing

It is widely agreed that language comprehension during reading is an incremental process in which a reader's interpretation of a text is constructed on a word-by-word basis (Pickering and Gambi, 2018, pp. 1002–1003). It had long been assumed that the primary mechanism driving language comprehension was integration (Dell and Chang, 2013; Ferreira and Chantavarin, 2018).[5] However, more recent approaches typically assume that language comprehension is an inherently Bayesian process, in which readers base their interpretation of current input on their expectations, leading them to predict forthcoming words (Ferreira and Chantavarin, 2018). Readers consequently do not delay in executing lexical access (i.e., visual world recognition), parsing (sentence processing) and semantic/discourse analysis (Pickering and Gambi, 2018, pp. 1002–1003). This sub-section aims to briefly discuss these stages, while at the same time acknowledging their empirically attested interaction (Taraban and McClelland, 1988; MacDonald, Pearlmutter and Seidenberg, 1994).

[5] Integration concerns "the linking of new ideas and concepts to what is already known or established" (Gernsbacher, 1991; Kintsch, 1991; Ferreira and Chantavarin, 2018).

The first stage of language comprehension (i.e., reading) involves lexical access, during which comprehenders visually identify word-related information (e.g., orthography, phonology, semantics). For example, when you see the string of letters, R A T, you think of a rodent with a pointed snout and a long tail. However, when this string is arranged in a different order, such as A R T, a piece of creative work comes to mind. The process of visual word identification (i.e., lexical access) is the starting point from which language comprehension can occur. Numerous computational models have been advanced to capture the cognitive mechanisms and processes behind visual word recognition, including the Interactive-Activation Model (McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1982), the Dual Route Cascaded Model (Colheart *et al.*, 2001), the Parallel Distributed Connectionist Model (Seidenberg and McClelland, 1989; Plaut *et al.*, 1996; Harm and Seidenberg, 2004) and the Connectionist Dual Process Plus (CDP+) Model (Perry, Ziegler and Zorzi, 2007). Although the general architecture differs between the individual models, each model assumes three different levels at which words are represented within the cognitive system: orthographic, phonological and semantic.

An interesting finding from this body of research is the influence of three specific word properties on lexical processing: frequency (Rayner and Duffy, 1986), length (Juhasz *et al.*, 2008) and predictability in context (Ehrlich and Rayner, 1981). This research has found that readers spend less time processing more frequently occurring words than less frequently occurring ones, as evidenced by studies from Rayner and colleagues (Inhoff and Rayner, 1986; Rayner and Duffy, 1986). Further research has shown that word length modulates the time course of lexical processing, with readers taking more time to process longer words (Juhasz, Inhoff and Rayner, 2005; White, Rayner and Liversedge, 2005; Juhasz, 2008; Juhasz *et al.*, 2008). Finally, there is evidence to suggest that highly predictable words are skipped more often, attract shorter fixation durations and are less likely to receive refixations following a regression in comparison to less predictable words (Balota, Pollatsek and Rayner, 1985; Schustack, Ehrlich and Rayner, 1987; Altarriba *et al.*, 1996; Rayner and Well, 1996).

Following visual word identification, readers engage in syntactic parsing. Parsing is considered to be "quick, automatic, and effortless", a process which we are typically not consciously aware of (Staub, 2015). At this stage, comprehenders must compute the grammatical structure of each sentence and establish the position of each word within this structure (Staub, 2015). Although it is widely agreed that language comprehension is a highly incremental process, cognitive psychologists remain divided as to whether the sentence processing system builds a single syntactic analysis and interpretation or can entertain multiple analyses at any one time when ambiguities arise. Such a question relates to whether the sentence processing system is serial or parallel, respectively. Serial models of syntactic parsing assume that the system only holds one analysis of a sentence at any one time; this analysis is then replaced in the event of an inaccurate interpretation. The seminal serial model is the garden path model (Frazier, 1978, 1987); more recent models include the ACT-R-based parser (Lewis and Vasishth, 2005) and the rational left-corner parser (Hale, 2011). Parallel models, however, assume that the system may hold several analyses at any one time, but that each analysis differs in their degree of activation. In other words, the processing system will construct multiple analyses of ambiguous content and will not choose between these analyses until faced with unequivocal disambiguating content. Such a view was first formalised by the constraint-based approach (MacDonald, Pearlmutter and Seidenberg, 1994; McRae, Spivey-Knowlton and Tanenhaus, 1998) and more recently by the surprisal model (Levy, 2008). The race model is a hybrid serial-parallel model; it predicts that in cases of ambiguity, multiple analyses are constructed, and that the comprehender selects the analysis that is the first to finish, meaning that only one complete analysis is considered (Traxler, Pickering and Clifton, 1998; Gompel, Pickering and Traxler, 2000, 2001).

In order to understand the intended meaning of a sentence or text, a comprehender must move beyond the basic syntactic relations of a sentence and perform semantic and/or discourse analysis. Semantic processing (i.e., interpretation) involves readers integrating the meaning of individual words in order to build a representation of the sentence's overall meaning, "determining what the various expressions in the sentence refer to, and who did what to whom" (Staub, 2015, p. 202). Discourse processing refers to the manner in which individuals derive meaning from larger, connected segments of text, such as activating information from memory and establishing connections within and between linguistic contexts (Graesser and Forsyth, 2013).

Similar to visual word recognition and syntactic parsing, various hypotheses have been advanced to explain how individuals extract the meaning of a text, including Construction-Integration Theory (Kintsch and van Dijk, 1978; Kintsch, 1991, 1998), the Structure Building Framework (Gernsbacher, 1991) and the Event Indexing Model (Zwaan and Radvansky, 1998). Despite a lack of agreement concerning the exact mechanisms behind this process, these models appear to agree that the successful comprehension of a text depends on both thematic and connecting processes in order to establish a precise, in-depth and coherent mental representation of the text.

2.2.2 L2 Sentence Processing

Most language processing research to date, however, has focused on typically developing, skilled monolingual readers. One aspect of human behaviour that significantly modulates language processing is the knowledge of more than one language. Recent years have witnessed a significant increase in the number of studies investigating the extent to which non-native speakers can apply grammatical knowledge in real-time sentence processing. A number of hypotheses have been advanced to account for fundamental L1/L2 differences, or even similarities, in real-time sentence processing. More specifically, these studies have sought to ascertain "whether L2 learners incrementally process the input (attempting the analysis of each word as it enters the current parse), and where misanalyses take place, whether learners can revise their interpretations in real time" (Roberts, 2019, p. 31). This section aims to discuss these theories in more detail. Although we focus predominately on language processing at the sentence and meaning level, we provide a brief discussion of L2 lexical processing.

2.2.2.1 Lexical Processing

Researchers interested in L2 lexical processing have explored two main themes in recent years: lexical entrenchment and cross-language activation. This sub-section section aims to provide a brief overview of these theories given the role that lexical properties play in the processing of the French subjunctive.

The Lexical Entrenchment Hypothesis (Diependaele, Lemhöfer and Brysbaert, 2013) predicts that word frequency effects are stronger in the L2 than in the L1, not because of qualitative differences in lexical processing, but rather due to differences in exposure to the target language. That is, speakers with higher levels of exposure to the target language are less likely to exhibit word frequency effects. For example, research has found that vocabulary size is a strong predictor of word frequency effects and that quantitative differences between L1 and L2 processing are significantly reduced, and sometimes non-existent, when L2 speakers have access to a wider vocabulary (Brysbaert, Lagrou and Stevens, 2017).

The Bilingual Interactive Activation Plus (BIA+) Model (Dijkstra and van Heuven, 2002), on the other hand, proposes that the cross-linguistic activation of lexical items impedes lexical processing; this is when orthographic, phonological and semantic features overlap between languages. Several studies have found that the effects of cross-linguistic activation are more observable at lower levels of proficiency, but gradually weaken as L2 speakers become more proficient (Libben and Titone, 2009; VanAssche *et al.*, 2011). Although these two themes have traditionally been discussed in parallel, recent research by Whitford and Titone (2019) suggests that lexical entrenchment and cross-language activation interact in order to influence bilingual word processing.

Of particular interest to this study is the accuracy and consistency of the lexical representation, that is, "the stability and accuracy of the language user's knowledge of a word's form, meaning, and use" (Perfetti and Hart, 2001; Perfetti, 2007; Kaan, 2014). Kaan (2014, p. 267) argues that the quality of L2 learners' lexical representations tend to be lower than native monolingual speakers, resulting in "a less consistent and more effortful retrieval of the information". This means that even though L2 speakers may demonstrate similar frequency biases to monolingual speakers, their lexical representations of the information may not be as consistent as native speakers, resulting in a reduced ability to predict upcoming linguistic information. Kaan (2014, p. 268) further proposes that L2 learners' ability to predict may develop as proficiency increases given that L2 learners are typically expected to develop higher quality lexical representations at more advanced levels of proficiency, "which can be retrieved consistently and easily enough to be used predictively". This claim has found empirical support in various studies demonstrating that predictive processing improves as a function of proficiency (Chambers and Cooke, 2009; Dussias *et al.*, 2013; Hopp, 2013; Kaan, 2014).

2.2.2.2 Sentence Processing

As previously mentioned, recent years have witnessed an emergence of studies investigating the processing of L2 grammatical knowledge in real-time. More specifically, and in a similar fashion to the theoretical debate in GenSLA, these studies have sought to investigate whether parsing differences are indicative of "qualitative or fundamental differences between L2 learners and native speakers" (Roberts, 2018, p. 36), or whether individual differences, such as the use of memory retrieval mechanisms, proficiency, processing speed or cognitive capacity, serve as a more accurate explanation for such differences (Dekydtspotter, Schwartz and Sprouse, 2006; Hopp, 2010; Sorace, 2011; Cunnings, 2017; Grüter, Rohde and Schafer, 2017). This sub-section aims to provide an overview of these theories.

A much-debated theory of L2 processing is the Shallow Structure Hypothesis (SSH, Clahsen and Felser, 2006, 2018). The SSH claims that even near-native L2 speakers have difficulties "building or manipulating abstract syntactic representations in real-time"; instead, L2 comprehenders may give preference to "semantic, pragmatics, or other types of information" and demonstrate heightened sensitivity to these types of information relative to L1 users (Clahsen and Felser, 2018, p. 3). The SSH therefore proposes gradual differences in L1/L2 processing, stating that "the representations adult L2 learners compute for comprehension are shallower and less detailed than those of native speakers […] and rely more on non-structural information in parsing" (Clahsen and Felser, 2006, 2018, p. 3). This does not mean that L2 users cannot utilise syntactic representations in real-time, but that they instead underexploit syntactic information and have trouble deploying grammatical knowledge during real-time processing. Numerous studies have tested this hypothesis, particularly in the context of morphosyntactic processing. However, findings have been mixed. For example, some studies have shown that L2 learners have difficulty manipulating grammatical information in order to accurately detect morphosyntactic violations during real-time processing (Love, Maas and Swinney, 2003; Jiang, 2004, 2007; Scherag, 2004; Marinis *et al.*, 2005a; Mueller *et al.*, 2005; Felser and Roberts, 2007; Sato and Felser, 2007; Keating, 2009), whereas other studies suggest that L2 learners are able to make use of grammatical information (Hopp, 2006; Jackson, 2008; Sagarra and Herschensohn, 2010; Tokowicz and Warren, 2010; Foote, 2011).

Other scholars, however, have argued against qualitative or fundamental differences between L2 learners and native speakers. Based on research examining syntactic ambiguity resolution, Cunnings (2017) highlights that both L1 and L2 processing patterns are characterised by variable attachment preferences and suggests that shallow processing may not necessarily account for L1/L2 differences, even though L2 are more likely to give preference to semantic- and discourse-level information during ambiguity resolution (e.g., Pan and Felser, 2011; Roberts and Felser, 2011; Pan, Schimke and Felser, 2015). Instead, these differences are said to result from how speakers encode and retrieve information from memory during sentence and discourse processing in light of previous studies (e.g., Roberts and Felser, 2011; Hopp, 2015; Jacob and Felser, 2016; Pozzan and Trueswell, 2016) showing that L2 learners experience "difficulties in erasing the initially assigned interpretation from memory" (Cunnings, 2017, p. 672). In particular, Cunnings maintains that L2 learners weigh discourse cues more heavily than other information types and must therefore negotiate increased competing elements when retrieving information from memory than native speakers. This consequently means that L2 learners are likely to exhibit increased sensitivity to interference compared with native speakers due to the different weightings of information types.

The Reduced Ability to Generate Expectations (RAGE) Hypothesis (Grüter, Rohde and Schafer, 2017), on the other hand, makes explicit predictions relating to expectancy-driven L2 processing. The authors propose that L1/L2 differences in processing result directly from L2 learners' limited ability, but not inability, to anticipate upcoming linguistic information. More specifically, L1/L2 differences are argued to be more evident in discourse-level predictions. Despite studies demonstrating L2 speakers' inability to actively anticipate upcoming information (Marinis *et al.*, 2005b; Dallas, 2008; Kaan, Dallas and Wijnen, 2010), other studies have found evidence of native-like predictive processing among L2 learners (Dussias *et al.*, 2013; Hopp, 2013). Kaan (2014) therefore argues that positing L1/L2 differences in (predictive) processing ability contradicts available evidence and cannot explain why L2 speakers can predict upcoming information in some contexts, but not in others. Nevertheless, a proposal that considers how prediction unfolds in L2 sentence processing is promising, especially in light of the recent focus on prediction in language processing, a topic that has been largely ignored by previous models of comprehension and integration processes (Ferreira and Chantavarin, 2018).

Further research in the field of L2 processing suggests that individual differences may account, at least in part, for the variability in L2 sentence processing. For example, limited capacity models (e.g., McDonald, 2006; Hopp, 2010; Sorace, 2011; Dekydtspotter and Renaud, 2014) couch their explanations for non-target-like L2 sentence processing in terms of the cognitive resource constraints (e.g., working memory restrictions) sub-serving syntactic processing. These models propose that L2 readers lack sufficient computational resources in order to integrate (linguistic) information successfully, "such that a full target-like parse cannot reliably be effected and shallow, incomplete or L1-based outputs are computed" (Hopp, 2017a, p. 123).

Some models argue that L1/L2 differences in feature configurations may account, at least in part, for the variation in L2 learners' sensitivity to morphosyntax. For example, if L1 properties do indeed modulate L2 processing profiles, structures with similar L1/L2 features should be expected to require comparatively less processing effort than those with different L1/L2 features (Tolentino and Tokowicz, 2011). Such a view has been formalised under the Morphological Congruency Hypothesis (Jiang *et al.*, 2011), according to which L1 properties modulate the L2 acquisition of morphosyntax, and the Unified Competition Model (MacWhinney, 2011), which attributes a determinist role to L1 influence in sentence interpretation. For example, MacWhinney proposes that when linguistic properties are expressed similarly in the L1 and L2, positive transfer obtains. These effects, however, may also be modulated by the weight of linguistic cues[6] available to express grammatical features in the L1 as well as L2 proficiency levels.

The Full Transfer/Full Access/Full Parse Hypothesis (Dekydtspotter, Schwartz and Sprouse, 2006) has also directly explored the role of the first language in real-time L2 processing. This hypothesis builds on the Full Transfer/Full Access Hypothesis (Schwartz and Sprouse, 1994, 1996) and assumes that the initial state of L2 development consists of the L1 grammar and makes specific predictions relating to parsing. In particular, the proponents of this hypothesis argue that L2 parsing must logically be deep (i.e., syntactically detailed) given the

---

[6] According to MacWhinney (2011, pp. 122–123), cue strength is influenced by the following four factors: task frequency, availability, simple reliability and conflict reliability.

dependent relationship between grammatical development and parsing, contra Clahsen and Felser (2006). According to the authors, target-deviant processing patterns result from biases introduced by the L1 grammar and parser, namely language-specific featural compositions.

2.2.3 Unification of Processing and Acquisition

Processing and acquisition have traditionally been treated as two independent constructs in L2 research, particularly within the generative paradigm. As a discipline, generative researchers have primarily focused on property-based theories of L2 acquisition. It is only in recent years that there has been increasing interest in transition theories of L2 acquisition, that is, how learners move from one state of linguistic knowledge to another. Such a question is not new to the field of language acquisition, however. In fact, it has been largely debated in the context of L1 acquisition from a variety of theoretical perspectives, including input-driven and knowledge-driven approaches. Input-driven approaches claim that learned knowledge is simply an imitation of the input to which the learner is exposed (Elman *et al.*, 1996). The acquired linguistic representations (the learned knowledge) is said to represent a compressed memory representation of the patterns detected in the input, with recently encountered input triggering previously encountered input by means of a similarity measure. Knowledge-driven approaches, however, postulate that the learner analyses the input looking for linguistic cues in order to establish these abstract representations. Aspects of these representations are considered to be innate and govern the cues that learners apply to construct them (Fodor, 1998a, 1998b; Lightfoot, 1999). The current sub-section focuses on both L1 and L2 development within a generative (i.e., knowledge-driven) paradigm.

2.2.3.1 L1 Theories

Fodor's (1998b, 1998a) 'Parsing to Learn' Hypothesis, also known as the 'Structural Triggers' Learning Model (Fodor and Sakas, 2017, p. 265), has been seminal in shaping how we understand learnability within a generative framework. It is based on the idea that mechanical parameter setting alone cannot explain how learning occurs: "Sentences must be parsed to discover the properties that select between parameter values" (Fodor, 1998b, p. 339). She argues that the human sentence processing system (the parser) plays a deterministic role in learning, such that a "special-purpose acquisition algorithm" is not required to explain learning (Fodor, 1998b, p. 341). Instead, the parser is argued to be fully innate and universal, which negates the need for language-specific parsing routines. As such, a learner should have access to the parametric values of the input grammar via UG prior to their initial exposure to the input. Exposure enables the learner to select, or set, the value required.

Under this view, parameter values are conceptualised as UG-sanctioned "treelets". A treelet represents a small syntactic tree that combines a set of underspecified syntactic nodes. An example of this would be a treelet involving a verb phrase node, with a verb and preposition phrase as its daughters. If a grammar includes such a treelet, the speaker could parse a sentence such as *Look at the dog*. However, if the grammar does not contain this treelet, then the parse of such a sentence would fail. Fodor argues that learners apply innate parametric "treelets" to rescue a parse of new input in cases of parsing failures that arise when the necessary parametric value is absent. These parsing failures in turn activate the need to generate new parameter values. Thus, it is not language-specific parsing routines *per se*, but rather different grammars sub-serving parsing that account for differences in sentence parsing routines.

What are the implications of this approach for language acquisition? In order to acquire new grammatical features, a learner must first parse the input by activating a supergrammar, or the best available grammar. This grammar typically consists of the current grammar and all the UG-constrained properties. If the learner's supergrammar is able to successfully parse the input, it is assumed that processing would occur in a similar fashion to how it would proceed with an adult native speaker. Conversely, in the case of processing failure, the learner is required to activate "the store of parametric treelets that UG makes available, seeking one that can bridge the gap in the parser tree" (Fodor and Sakas, 2017, p. 266). If we assume that learning is an incremental process, then with every successful parse using a new treelet, the activation threshold of the treelet would increase. Such a process would then repeat until the treelet is incorporated in the grammar.

2.2.3.2 L2 Theories

In recent years, several unified frameworks integrating both representational and processing accounts of second language acquisition have been proposed within a generative perspective. These frameworks can be largely divided into two categories: failure-driven (Carroll, 1999, 2001, 2007; Dekydtspotter and Renaud, 2014; Phillips and Ehrenhofer, 2015) and acquisition-by-processing (Sharwood Smith and Truscott, 2014) approaches. Due to space constraints, this sub-section focuses specifically on the Parser as the Language Acquisition Device (PLAD) as an example of a failure-driven approach.

Based on the notion of failure-driven learning, in particular Fodor's Parsing to Learn Hypothesis, Dekydtspotter and Renaud (2014) propose the parser as the language acquisition device. A fundamental assumption underlying this proposal is the idea that sentence processing is plausibly constrained by "a parser that generates UG-sanctioned representations that must be licensed by language-specific information at each step" (Dekydtspotter and Renaud, 2014, p.

133). This language-specific information is formalised by means of feature bundles for functional categories and lexical items. In this framework, parsing routines are said to be the by-product of the language-specific grammatical representations encoded in the functional lexicon.

With this in mind, Dekydtspotter and Renaud argue that the parser acts as the language acquisition device, relying on both the "incremental generation of UG-sanctioned representations" and the "licensing of these representations by an abstract parameterized lexicon" (Dekydtspotter and Renaud, 2014, p. 155). This means that if the learner is unable to parse a specific structure, the licensing conditions should, in theory, trigger the creation of new feature bundles containing functional morphemes and other lexical items, which would result in the development of L2 linguistic knowledge.

Put simply, Dekydtspotter and Renaud build their proposal on the fundamental assumption that the incremental parsing of input is the driving force behind grammatical learning (Fodor, 1998b, 1998a). Parsing failures impel the learner to apply the supergrammar and access the most economical structure in order to parse the input sentence. Assuming that the parser continuously monitors the success or failure of each structure parsed, then it should follow that the most successfully parsed structures will be activated more frequently up until the point where that structure is incorporated into the grammar. As a result, such a process does not require an additional learning mechanism. Under this view, the L2 grammar acquisition mechanism represents the "parser embedded in the processing system" (Dekydtspotter and Renaud, 2014, p. 156).

However, as we already know, acquisition is incremental and involves periods of stability and change as learners access new feature bundles. How can we explain these stages? The process of licensing can account, in large part, for the incremental nature of learning. Licensing requires the learner to retrieve the relevant grammatical specifications from a memory store, with working memory placing restrictions on processing, such that grammatical specifications must be accessed within a somewhat narrow timeframe. As discussed previously, limited capacity models (e.g., McDonald, 2006; Hopp, 2010; Sorace, 2011; Dekydtspotter and Renaud, 2014) couch their explanations for non-target-like L2 sentence processing in terms of cognitive resource constraints (e.g., working memory restrictions) sub-serving processing. Stages in acquisition therefore result from the constantly evolving relative strengths of activation, which arise from both licensing failures and successes during parsing. It thus follows that learners' grammar experience "a period of transition in which two values may be in competition" (Dekydtspotter and Renaud, 2014, p. 157).

In summary, Dekydtspotter and Renaud (2014, p. 131) propose that "the PLAD instantiates feature re-assembly (Lardiere, 2009) in response to licensing failures (Clark and Roberts, 1993), characterizing transitions between grammatical states (Gregg, 1996, 2003) in the Full Transfer/Full Access model (Schwartz and Sprouse, 1994, 1996)."

2.2.4 Eye-Tracking in Sentence Processing Research

In order to gain a deeper understanding of the cognitive processes underlying sentence processing, psycholinguists have employed a number of experimental techniques, including self-paced reading and eye-tracking during reading. The current section aims to introduce the reader to eye-tracking as an experimental paradigm for sentence processing research.

During reading, our eyes appear to progress smoothly through the text for the most part, with only minor disruptions to reading, for example, when we move to the next line of a text or refer back to an earlier word in the sentence (Rayner, 2009). The reality, however, is that eye movement is far from smooth; it is typically characterised by a series of fixations and saccades. Fixations are the intervals of time (200-250 ms on average) when the eyes remain relatively still, whereas saccades refer to the periods of time (typically 30-40 ms) when the eyes are moving (Rayner, 2009).

Saccades, however, do not always involve forward movement. Regressions (i.e., backward movement) occur approximately 10-15% of the time when the eyes move back to previously encountered parts of the text (Rayner, 1998). The duration of these regressions can vary. Short regressions typically occur when a reader overshoots the target. Longer regressions, on the other hand, are associated with the difficulty of the text (Rayner *et al.*, 2004).

The cognitive system typically only acquires (i.e., perceives and processes) new visual information during fixations; it is during these fixations that the cognitive system plans when and where to move the eyes next (Rayner, 2009). During saccades, however, eye movement is thought to be so quick that the brain suppresses the processing of new visual input but can still continue processing already perceived information (Matin, 1974; Irwin and Carlson-Radvansky, 1996; Irwin, 1998). This is known as saccadic suppression (Matin, 1974).

Researchers interested in eye movement data have long debated whether there is an eye-mind link and more recently how strong it is. The concept of the eye-mind link refers to the notion that during complex processing tasks, such as reading, the location of the eyes (i.e., fixation patterns and, to a lesser extent, the direction of saccades) provides an indication of attention and thus online cognitive processing (Liversedge and Findlay, 2000; Rayner, 2009). In other words, a reader's eyes can tell us what they are focusing on and the level of cognitive effort exerted in order to process the input at the point of fixation. Factors, such as the difficulty and complexity of the visual input, have the potential to influence fixations and saccades. For example, when we read a syntactic or semantic anomaly, our eyes typically fixate on the anomaly for longer than normal (Rayner *et al.*, 2004). In order to analyse eye movement data, researchers have devised a number of reading time measures. These measures are described in the following section.

In recent years, the focus has shifted away from whether there is an eye-mind link to how strong the eye-mind link is and how it should be conceived (Godfroid, 2019). Research has shown that eye movement behaviours appear to reflect oculomotor phenomena that arise, irrespective of linguistic and other cognitive processes (Nuthmann, Engbert and Kliegl, 2007; Henderson and Luke, 2012). In contrast, other studies have found that eye movements often index higher cognitive processes and are modulated by a combination of linguistic factors, including word predictability, frequency and syntactic complexity (Rayner, 1998, 2009).

In the early years, the debate was populated by two competing positions: the oculomotor model, which, at least in its strongest form, did not posit any immediate role for cognitive factors in modulating individual eye movements (e.g., O'Regan and Levy-Schoen, 1987; O'Regan, 1990, 1992), and the processing model, which, again in its strongest form, argued that ongoing cognitive processing has complete moment-by-moment control of eye movements (e.g., Just and Carpenter, 1980; Morrison, 1984). Our discussion relating to the distinction between serial and parallel processing during processing in Section 2.2.1 can be applied to both cognitive control and oculomotor models. Typically, parallel processing is more strongly linked to oculomotor models and serial processing to cognitive control models. Recent years, however, have given way to a more nuanced debate exploring the relative involvement of low-level oculomotor factors and higher-level cognitive factors with eye movement control (Starr and Rayner, 2001).

Several computational models have been proposed in order to capture these differences. These models have been shown to accurately predict eye movement patterns based on specific properties of the text. The two most eminent models are the serial-attention E-Z reader (Reichle *et al.*, 1998; Reichle, Pollatsek and Rayner, 2012) and the attention-gradient SWIFT model (Engbert *et al.*, 2005). It must be noted that these computational models have been almost exclusively based on skilled, monolingual adult readers. Unfortunately, however, it is beyond the scope of this thesis to discuss these models in any further detail.

Nevertheless, the ability to track and measure a reader's eye movement provides researchers with high spatial-temporal resolution information about the moment-to-moment linguistic and cognitive processes behind language comprehension. By recording an individual's eye movement during silent reading (i.e., as participants read stimuli on a computer screen), we can accurately identify the exact location that any disruption to language comprehension occurs and the subsequent impact it has on the time course of language processing. While techniques, such as self-paced reading, provide similar information to a certain extent, it is predominately restricted to earlier measures of language processing, such as first (pass) reading time, which indicate sensitivity to syntactic and/or semantic anomalies, and thus does not provide later measures, such as go-past time, regression rate, or rereading, which can tell us about the time spent resolving a problem or difficulties in integration and/or higher-level processing (e.g., discourse processing).

Eye movement data is particularly beneficial to the study of L2 populations. Research in the GenSLA paradigm has historically sought to tap into learners' underlying representations of language through the elicitation of offline comprehension data (i.e., form-focused tasks such as grammaticality judgments). However, mounting evidence would appear to suggest that traditional elicitation techniques such as these rely predominately on automatized explicit knowledge (Suzuki, 2017), rather than the more implicit underlying knowledge of language that GenSLA researchers are interested in. This can be particularly problematic when studying structures that are explicitly taught in the L2 classroom and that L2 learners have extensive metalinguistic knowledge of (Keating and Jegerski, 2015; Conklin and Pellicer-Sánchez, 2016). In theory, then, eye movement data should allow researchers to more directly tap into learners' unconscious, underlying knowledge of language, as it does not require additional artificial procedures. In this sense, eye movement data can be considered representative of both underlying linguistic representations and processing.

Eye movement data has the further advantage of being more ecologically valid than traditionally used psycholinguistic techniques, in that it does not interfere with natural reading. For example, most eye-tracking studies do not require the use of artificial procedures (e.g., where one word at a time is presented with no visual input visible in the parafovea) and tasks (e.g., where participants must press a button when an anomaly is detected) to identify the location and time course of difficulties in language comprehension.

2.2.4.1 Reading Measures

As discussed in the previous section, researchers have developed a number of reading time measures in order to analyse eye movement data. Although it is beyond the scope of this thesis to discuss every measure, Table 2.1 presents a selective overview of the key reading time measures, all of which will be analysed in the studies presented in this thesis. A combination of both early and late measures was specifically chosen in order to investigate how the experimental manipulations disrupted different stages of language processing.

Table 2.1 Reading measures to be analysed

Measure	Description	Explanation
Skipping probability	The likelihood that the region of interest is skipped on first pass reading.	Reflects lexical properties (e.g., predictability, length, frequency, etc.) of region of interest (Rayner, 2009).
First fixation duration	The duration of the first fixation on a region of interest.	Reflects lexical properties (e.g., predictability, length, frequency, etc.) of region of interest (Liversedge, Paterson and Pickering, 1998).
Gaze duration	The sum of all fixations on a region of interest before progressing to another word.	Reflects sensitivity to semantic and syntactic anomalies (Rayner *et al.*, 2004).
Go-past time	The sum of all fixations from the first time the eyes enter the region until leaving it to the right.	Can indicate lexical and integration difficulties leading to regression to earlier points (Rayner *et al.*, 2004).
Total reading time	The sum of all fixations on a region of interest.	Influenced by both early and late processing; significant effect of 'total reading time', but not of early measures, suggests a late effect on processing (Liversedge, Paterson and Pickering, 1998).
Regression-out probability	The percentage of times a regression was launched from the current region of interest.	Reflects higher-level processing (e.g., discourse processing) of a text (Rayner and Pollatsek, 1989).
Regression-in probability	The percentage of times a regression was launched from later interest areas to the current region of interest area.	Reflects higher-level processing (e.g., discourse processing) of a text (Rayner and Pollatsek, 1989).

2.2.4.2 L2 Eye Tracking Studies

Recent years have witnessed a notable increase in the number of studies that have used eye-tracking during reading to address questions relating to "the representation, processing and acquisition of grammar", many of which have predominately been couched within a formal linguistics framework (Godfroid, 2019, p. 65). In grammar-focused studies, researchers have been particularly interested in a sentence or a few sentences rather than a longer text. These studies can be divided into at least three paradigms. In the anomaly detection or violation paradigm, sentences include a grammatical, semantic, pragmatic, or discourse anomaly (Keating, 2009; Clahsen *et al.*, 2013; Sagarra and Ellis, 2013; Ellis *et al.*, 2014; Godfroid *et al.*, 2015; Lim and Christianson, 2015; Zufferey *et al.*, 2015; Hopp and León Arriaga, 2016). In the ambiguity resolution paradigm, sentences involve a syntactic ambiguity, such as a garden-path sentence (e.g., Dussias and Sagarra, 2007; Roberts, Gullberg and Indefrey, 2008; Chamorro, Sorace and Sturt, 2016). In the dependency paradigm, sentences consist of long-distance syntactic dependencies, such as *wh*-questions or relative clauses (Felser, Sato and Bertenshaw, 2009; Felser and Cunnings, 2012; Felser *et al.*, 2012; Boxell and Felser, 2017). The current sub-section aims to provide a selective overview of these studies, with a particular focus on the anomaly detection paradigm.

A central question within this paradigm is whether learners have acquired a specific grammatical structure with a particular focus on morphosyntax (e.g., tense, person, number, gender or case marking agreement). In order to examine learners' internal grammatical representations, researchers have operationalised acquisition as sensitivity to grammatical and ungrammatical stimuli (Godfroid, 2019). A core assumption underlying this approach is the idea that when the internal grammatical representations have been acquired, learners will demonstrate specific reading patterns (e.g., a slowdown in reading or some other reaction) in response to stimuli that violate their internal grammar.

One such study is Lim and Christianson (2015), who investigated sensitivity to subject-verb agreement violations and how task (reading for comprehension vs. translation) and proficiency modulate this sensitivity. The study involved a 2 x 2 design, manipulating the factors, grammaticality (grammatical G or ungrammatical UG) and match (whether a head and a local intervening noun matched in number (M) or not (No M)). Examples of test items are provided in (1) below.

(1) a. The teachers who instructed the students were very strict. (G, M)
b. The teachers who instructed the student were very strict. (G, No M)
c. *The teacher who instructed the student were very strict. (UG, M)
d. *The teacher who instructed the students were very strict. (UG, No M)

Findings from this study revealed that L2 sensitivity to subject-verb agreement depended on both L2 proficiency and task. In particular, the authors noted that although L2 speakers demonstrated delayed sensitivity to subject-verb agreement violations in the comprehension task, sensitivity to violations was greater in the translation task, with more rapid detection of violations as proficiency increased. These results thus suggested that L2 processing was influenced by reading goals and proficiency. More generally, the authors interpreted their findings as reflecting quantitative (i.e., strategically "good enough") rather than qualitative L1/L2 differences in processing.

Hopp and León Arriaga (2016) examined sensitivity to structural and inherent case marking violations by 24 Spanish L1 speakers and 27 German L2 speakers of Spanish. The structural cases violations involved objects of ditransitive verbs and inherent case violations involved objects of transitive verbs (differential object marking; DOM). Eye movement data revealed two key findings: (1) both L1 and L2 speakers exhibited sensitivity to unmarked objects in sentences with ditransitive verbs and (2) L1 speakers, but not L2 speakers, demonstrated longer reading times for unmarked animate and marked inanimate objects in transitive sentences. The authors argued that the non-target-like processing of differential object marking did not necessarily indicate incomplete acquisition or feature assembly. Instead, it reflected the fact that the L2 parser licensed underspecified interpretable features in real-time comprehension, yet instantly identified feature clashes.

Sagarra and Ellis (2013) tested sensitivity to tense incongruences by 61 low and 59 high proficiency learners of Spanish whose L1 was either English, Romanian or Spanish. Eye movement data revealed showed that all learners exhibited sensitivity to verb-adverb and adverb-verb inconsistencies. However, participants' reading times were longer at their preferred cue (i.e., the verb for morphologically rich languages or adverb for morphologically poor languages), regardless of its location. The only exception, however, was L1 speakers of morphologically rich languages (i.e., Romanian and Spanish) who spent longer processing the verb in incorrect sentences where the verb occurred before rather than after the adverb. Furthermore, previous L1 experience, in particular morphological richness, modulated fixation time on verb cues, but not adverbs. Finally, the study found that reading times and regressions were longer for low-proficiency learners than high-proficiency learners. The authors concluded that these findings supported theories of transfer effects, but did not necessarily challenge Clahsen and Felser's (2006) Shallow Structure Hypothesis since the target structure involved morphological, not syntactic, agreement.

Another study by Clahsen, Balkhair, Schutter and Cunnings (2013) explored the time course of processing morphologically complex words by 24 English L1 speakers and 25 Dutch L2 speakers of English. Eye movement data revealed that while category restrictions (against plurals inside derived words) modulated L2 online processing, morphological restrictions (against regular plurals inside derived words) did not. The authors concluded that the L2 processing of morphologically complex words was not only slower than L1 processing, but that the L2 parser used real-time grammatical analysis (e.g., morphological information) less frequently than the L1 parser.

Keating (2009) studied sensitivity to gender agreement violations in L2 Spanish by 18 L1 Spanish speakers and 18 beginner, 14 intermediate and 12 advanced L2 learners of Spanish. Eye movement data showed that although advanced Spanish learners were sensitive to local gender agreement, beginner and intermediate learners were. Furthermore, the study revealed that the distance between noun and adjectives influenced advanced L2 learners' ability to process gender agreements. Keating thus concluded that these findings supported Clahsen and Felser's (2006) Shallow Structure Hypothesis with respect to morphological processing.

To summarise, the studies presented above revealed an asymmetry between L1 and L2 parsing routines, as evidenced by eye movement data. These differences appear to be influenced by multiple factors, including proficiency, reading goals, transfer effects and type of linguistic cues. Overall, however, it would appear that these asymmetries suggest quantitative rather than qualitative differences in processing.

2.3 Conclusion

As discussed in the introduction, the current thesis brings together a theoretical understanding of both L2 acquisition and processing, integrating many of the theories discussed above, in order to investigate the parallel between acquisition and processing. In particular, it seeks to develop a greater understanding of how L2 sensitivity is influenced by a multitude of factors, including L1/L2 differences in feature configurations, proficiency, exposure via residence abroad and reading goals. A significant contribution of the current thesis is its experimental design, in particular the use of a more sensitive, implicit measure of linguistic knowledge (that is, eye-tracking during reading).

The subsequent study on the English subjunctive is motivated by the observation that the L1 feature configurations can modulate successful L2 acquisition, as per Lardiere's (2009) Feature Reassembly Hypothesis. Exploring the distributional and interpretational properties of the English subjunctive allows us to accurately establish the contexts in which native speakers of British English license the subjunctive in order to generate precise predictions about the role of L1 influence in the L2 development of the French subjunctive.

Chapter 3 Study 0: English Subjunctive

3.1 Introduction

It is widely accepted that Romance languages, such as French, Spanish and Italian, express the subjunctive mood.[7] However, the status of the subjunctive in languages such as English is widely debated. In fact, Denison (1998, p. 263) goes as far to say that "[i]n Br[itish] E[nglish] the present subjunctive [...] has retreated to high-flown literary or legal language". However, some studies, such as Övergaard (1995) and Leech *et al.* (2009), have argued that the twentieth century witnessed an increase in the use of the subjunctive form in subordinate clauses following mandative predicates,[8] such as in (2).

(2) They suggested that he be reprimanded.
(3) They suggested that he should be reprimanded.

(Crawford, 2009, p. 257)

Many of these studies have, in large part, shown that subjunctive use is more widespread in American English than in British English, with the former spearheading the rise in subjunctive use (e.g., Övergaard, 1995; Hundt, 1998, 2018). More specifically, this research has found that British English speakers typically prefer the modal construction in (3) as opposed to subjunctive complements in (2) (Crawford, 2009).

The present chapter aims to explore the availability of the subjunctive in British English further by presenting findings from an experimental study. The primary aim of this study was to establish the contexts in which British English allows the subjunctive in order to formulate precise predictions relating to the role of L1 influence on the development of the French subjunctive by British English-speaking L2 learners of French.

The current chapter is structured as follows. First, Section 3.2 describes subjunctive use in British English. Section 3.3 then outlines the research questions and methodology, focusing in particular on the experimental design and research participants, with Section 3.4 presenting the results. Finally, Section 3.5 discusses these results and their implications for L2 acquisition.

---

[7] Following Portner (2018, p. 4), we assume that mood represents "an aspect of linguistic form which indicates how a proposition is used in the expression of a modal meaning".

[8] Crawford (2009, p. 259) defines mandative clauses as "any clause in its finite verb form that explicitly addresses the fact that some person or entity wants a particular action to be taken or a certain event to happen".

3.2 English Subjunctive

As we have briefly discussed above, the English subjunctive is typically found in mandative contexts, but is not as widespread in English, particularly in British English, as it is in Romance languages. Examples of subjunctive triggers include *demand, order, propose, recommend, request* and *require*. In English, the subjunctive is marked by the bare (non-inflected) verb form (cf., Roberts, 1993), as in (4)-(6). However, this has not always been the case. Up until Middle English, subjunctive forms had distinct morphological endings (cf., Roberts, 1993).

(4) The professor demands that she be informed of the student's results.
(5) It is essential that every student enrolling in the course have the necessary qualifications.
(6) We recommend that the student not leave the classroom.

A distinguishing syntactic feature of the English subjunctive is the absence of *do*-support in negated sentences, such as in (6). Several scholars have postulated, based on the absence of *do*-support and the ungrammaticality of modal verbs with the subjunctive, that subjunctive clauses involve a functional projection headed by a morphologically independent zero mood marker (i.e., a null morpheme) (Roberts, 1985, 1993; Lasnik, 1995; Potsdam, 1996, *inter alia*).

Assuming that the English subjunctive form carries a zero morpheme, it could be argued that English subjunctive clauses share the same underlying structure as subjunctive clauses in French, where MoodP dominates Tense Phrase (TP). The question remains, however, whether *that*-subjunctive clauses project TP, given the absence of tense/aspect distinctions in English subjunctive clauses and the subsequent ungrammaticality of the subjunctive with aspectual verbs, as in (9).

(7) They had previously demanded that he not be executed yesterday.
(8) They are now demanding that he not be executed tomorrow.
(9) * They demand(ed) that he have been executed by midnight.

Examples taken from Radford (2007, p. 87)

Radford (2007) proposes that English subjunctive clauses lack a TP projection, given the position of *be* below *not* in (7)–(8). This proposal assumes that only T(ense) can trigger the movement of *be* to a position above *not*, as is the case with indicative clauses (e.g., *They are not* *accepting this offer*). Under this analysis, mandative subjunctive clauses in English would project the following structure:

(10)

$$\ldots V_M [_{CP} [_{ForceP}\ Force_{[uM]} [_{FinP}\ Fin_{[uM]} [_{IP}\ (DP) [_{MoodP} [_M \varnothing] [_{TP} \ldots]]]]]]$$

selection — checking (agree)

(Iverson, Kempchinsky and Rothman, 2008, p. 140)

The structure presented in (10) implicates three functional projections: the Force[9] and Fin[10] heads in the left periphery[11] as well as a Mood head in the inflectional field (Kempchinsky, 2009). Subjunctive use is said to result from a shift in the model of evaluation from the default model to the bouletic[12] model of the matrix subject (Quer, 2009). This shift is expressed syntactically using a modal feature, M, in the Force head of the left periphery, which in the case of mandative subjunctives is an uninterpretable feature. In line with Bianchi (2001), FinP is thought to represent the projection that indexes with the discourse context. Yet, in subjunctive clauses, FinP adopts the uninterpretable modal feature from Force; the deictic centre is situated within the bouletic intensional model and is only indirectly associated with the extensional speech context. The interpretable feature in the head of MoodP values and deletes the uninterpretable modal features in the left periphery. For a more in-depth discussion, see Section 4.2.

3.2.1 *For-to* infinitive as an Alternative Marker of Modality

It has long been acknowledged in generative research on sentential complementation that *for-to* infinitive complements are used much more frequently that *that*-subjunctives as complements to both emotive-factive and volitional predicates (Emonds, 1970, 1976; Bresnan, 1972).

(11) I really prefer for Lewis to be the main presenter.

(12) What I suggest is for Lewis to be the main presenter.

---

9 Fin describes the finite or non-finite nature of the clause (Rizzi and Bocci, 2017).

10 Force describes the illocutionary force or clause type (e.g., declarative, interrogative, exclamative, etc.). That is, the type of information that has to be available to a higher governor in case of embedded clauses (e.g., a main verb like *believe* would require a declarative, *ask* an interrogative, etc.) (Rizzi and Bocci, 2017)

11 The left periphery refers to the initial structural position of the clause, "consisting of sequences of hierarchically organized functional elements" (Rizzi and Bocci, 2017).

12 Bouletic modality relates to "what is possible or necessary, given a person's desires" (Fintel, 2006, p. 2)

According to Edmonds (1970, 1976), *for* acts as a type of subjunctive complementiser in English. In examples such as (13), we typically assume that it is phonologically null (i.e., not pronounced) given its presence in the clefted version of (14).

(13) The professors want $[_{CP}$ $[_{C}$ Ø] [James to present the paper]]

(14) What the professors want is for James to present the paper.

(15) * What the professors want is James to present the paper.

Van Gelderen (2004) suggests that the core interpretation of *for* in this context is [+future] with a "purposive meaning", as illustrated in (16). The use of *for* as an infinitival complementiser is therefore not compatible with epistemic predicates, as exemplified in (17)–(18).

(16) I would like for you to eat a Kit Kat.

(17) * I believe for her to be a spy.

(18) * What I really believe is for her to be a spy.

Iverson, Kempchinsky and Rothman (2008) suggest that a more accurate analysis of the infinitival complementiser *for* would be as a marker of intensional modality, where it merges (i.e., joins the syntax) as the head of FinP in the left periphery, as schematised in (19).

(19) $\ldots V_{M}\ [_{CP}\ [_{ForceP}\ Force_{[uM]}\ [_{FinP}\ [_{Fin}\ \textit{for}_{M}][_{TP} DP\ [_{T'}\ \textit{to}\ VP]]]]$

selection checking

In other words, *for* appears to serve as a morphosyntactic form which detects the uninterpretable modality feature in Force. However, if we follow the narrow definition of mood as a set of inflectional morphemes on the verb, then it would be misrepresentative to refer to the *for-to* variant as the subjunctive mood.

3.2.2 Dialectal Differences in Subjunctive Use

As we discussed briefly in the introduction, several studies have reported dialectical differences in subjunctive use between British and American English, noting that while American English favours the subjunctive form, British English prefers the modal construction in mandative clauses.

One of the first studies to examine this, using data from the Brown and LOB corpora and a group of pre-determined mandative predicates, observed that the subjunctive form was more frequently used in American than in British English (Johansson and Norheim, 1988). Several follow-up studies, using an identical group of triggers and data from the Brown family of corpora, have found (1) that the rise in subjunctive use in American English occurred predominately in the first part of the twentieth century and (2) that the subjunctive in mandative clauses in British English started to increase in the second half of the twentieth century (Hundt, 1998; Hundt and Gardner, 2017; Waller, 2017). However, it would appear that the rate of change in British English is much slower than in American English, to the extent that this change may already be easing at a lower level (Hundt, 2018).

It is clear from the discussion pursued above that the status of the mandative subjunctive, particularly in British English, is open to debate. However, in order to adopt a Feature Reassembly approach to the L2 acquisition of the French subjunctive, it is essential that we accurately define the licensing contexts for the subjunctive in both English and French. With this in mind, the current study aims to investigate the licensing contexts for the subjunctive in British English.

3.3 Method

In this experiment, we examined the availability of grammatical mood in the complement clauses of verbs spanning three lexical-semantic categories: desire, directive and emotive-factive. Initially, we anticipated that British English speakers would accept indicative CPs with directive and emotive-factive verbs, *for-to* infinitives with desire and emotive-factive verbs and subjunctive CPs with directive verbs. The primary aim of this experiment therefore was to test these predictions and identify the contexts in which British English speakers accept the subjunctive, with a view to establish precise predictions about the role of L1 influence in the L2 development of the French subjunctive. In particular, we sought to address the following research question:

- Do L1 speakers of British English license the subjunctive in complement clauses?
- If so, in which contexts do these speakers accept the subjunctive in complement clauses?

All participants completed the experiment via iSurvey, an online survey generation tool available to members of the University of Southampton. The experiment involved two tasks. The first was a short background information questionnaire, involving questions about age, gender, education level, city and country of origin, first language and knowledge of any additional languages. The second was an (untimed) grammaticality judgment task. In this task, participants were asked to read each sentence silently and then judge how acceptable each sentence sounded by selecting a number between "1" and "7", with "1" indicating an entirely unacceptable sentence and "7" a perfectly acceptable sentence.

The experiment consisted of 54 target items in total, split evenly across 9 conditions. A further 16 filler items were also included, such that participants saw a total of 70 test items. We manipulated the following factors: structure type (*for-to* infinitives, indicative CPs and subjunctive CPs) and matrix semantic property (desire, directive and emotive-factive). Examples of the conditions are provided in Table 3.1. The presentation order of the test items was randomised for each participant.

Table 3.1 Examples of test items

Semantics	Structure	Example
desire	indicative	Cécilia wants that her friend has more confidence in her (abilities).
	subjunctive	Cécilia wants that her friend have more confidence in her (abilities).
	(for)-to infinitive	Cécilia wants her friend to have more confidence in her (abilities).
directive	indicative	The minister demands that the conference takes place before the end of the year.
	subjunctive	The minister demands that the conference take place before the end of the year.
	(for)-to infinitive	The minister demands the conference to take place before the end of the year.
emotive-factive	indicative	The dentist likes that his patient is aware of the risks associated with eating sugar.
	subjunctive	The dentist likes that his patient be aware of the risks associated with eating sugar.
	(for)-to infinitive	The dentist likes his patient to be aware of the risks associated with eating sugar.

3.3.1 Participants

Participants included 18 university-level educated first language speakers of British English (10 females; 8 males) with an average age of 25.28 (range 19–52). Most participants reported knowledge of at least one additional language, including Chinese, French, German and Spanish.

3.4 Results

The acceptability judgment data was analysed using a mixed-effects linear regression model, fitted with the *lmer* function of the *lme4* package (Bates *et al.*, 2015) in the R environment (R Development Core Team, 2014). The fixed effects structure included verb type (desire, directive, emotive-factive) and structure type (*for*, indicative, subjunctive) as well as the fixed interaction term, verb type × structure type. The (categorical) fixed effects were sum coded. For verb type, we examined the difference between desire and directive sentences (directive –1, desire 1, emotive-factive 0) and between emotive-factive and directive sentences (directive –1, desire 0, emotive-factive 1). For structure type, we explored the difference between the *for* structure and indicative CPs (*for* 1, indicative –1, subjunctive 0) and between subjunctive CPs and indicative CPs (for 0, indicative -1, subjunctive 1). Initially, we attempted to fit the model using the 'maximal' random effects structure (i.e., slopes for each fixed effect across item and subject) that converged (Barr *et al.*, 2013). However, this had to be cut down due to non-convergence or a singular fit (i.e., perfect or near perfect correlations in the random structure). Absolute t values exceeding ± 1.96 were analysed as significant.

Table 3.2 Mixed-effects models for grammaticality judgment task

Fixed Effect	*b*	*SE*	*t*
(Intercept)	4.39	0.21	20.47
Semantic: Desire vs. Directive	-0.11	0.27	-0.41
Semantic: Emotive-Factive vs. Directive	-0.58	0.27	-2.14
Structure: For vs. Indicative	0.13	0.08	1.65
Structure: Subjunctive vs. Indicative	-0.87	0.08	-11.33
Semantic (Desire vs. Directive) × Structure (For vs. Indicative)	2.05	0.11	18.89
Semantic (Emotive-Factive vs. Directive) × Structure (For vs. Indicative)	-0.73	0.11	-6.78
Semantic (Desire vs. Directive) × Structure (Subjunctive vs. Indicative)	-0.31	0.11	-2.89
Semantic (Emotive-Factive vs. Directive) × Structure (Subjunctive vs. Indicative)	-1.05	0.11	-9.72

Note: Shaded cells denote statistical significance ($p < 0.05$).

The model parameters are presented in Table 3.2. We found a significant main effect of verb type (but only for emotive-factive vs. directive) and of structure type (but only for indicative vs. subjunctive). All interaction terms included in the model were significant: semantic (desire vs. directive) × structure (for vs. indicative), semantic (emotive-factive vs. directive) × structure (for vs. indicative), semantic (desire vs. directive) × structure (subjunctive vs. indicative), semantic (emotive-factive vs. directive) × structure (subjunctive vs. indicative).

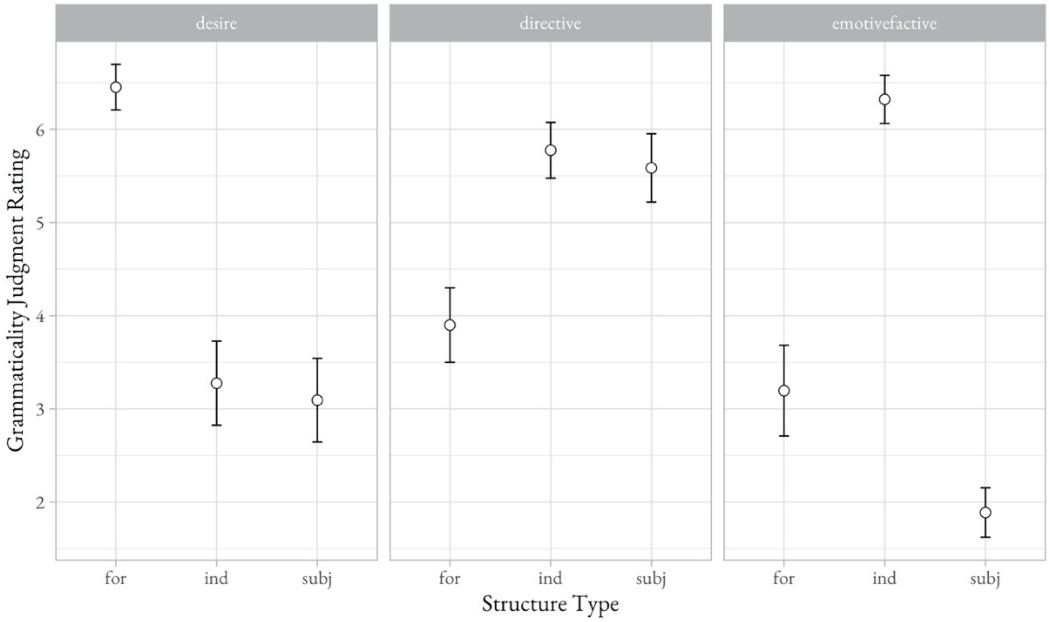


Figure 3.1 Mean grammaticality judgment ratings (with standard errors bars) for each structure, grouped according to verb type

To explore all the contrasts that were present in the two-way interactions, we computed a series of pairwise comparisons using the *emmeans* package (Lenth, 2018) in the R environment (see Table 3.3) and plotted the mean acceptability ratings and standard errors for each structure and verb type (see Figure 3.1). Together, these outputs revealed that native speakers of British English demonstrated a distinct preference for the *for-to* structure with desire verbs, rating both the indicative- and subjunctive CP structures as unacceptable. However, speakers found the indicative- and subjunctive CPs structures equally acceptable with directive verbs but considered the *for* structure (borderline) unacceptable. With emotive-factive verbs, it appears that out of the three possible structures, only the indicative CP structure was acceptable for British English speakers.

Table 3.3 Pairwise comparisons for grammaticality judgment task

	Contrast	*b*	*SE*	*df*	*t*	*p*
Desire	For vs. Indicative	3.16	0.23	923.10	13.76	<.0001
	For vs. Subjunctive	3.35	0.23	923.00	14.61	<.0001
	Indicative vs. Subjunctive	0.19	0.23	923.00	0.82	1.00
Directive	For vs. Indicative	-1.87	0.23	923.10	-8.12	<.0001
	For vs. Subjunctive	-1.69	0.23	923.00	-7.36	<.0001
	Indicative vs. Subjunctive	0.19	0.23	923.10	0.81	1.00
Emotive-factive	For vs. Indicative	-3.14	0.23	923.00	-13.67	<.0001
	For vs. Subjunctive	1.31	0.23	923.10	5.71	<.0001
	Indicative vs. Subjunctive	4.45	0.23	923.10	19.30	<.0001

3.5 Discussion and Conclusion

The current study set out to investigate the contexts in which British English speakers allow the subjunctive in their L1 grammar. Initially, we anticipated that these speakers would accept indicative CPs with directive and emotive-factive verbs, *for-to* infinitives with desire and emotive-factive verbs and subjunctive CPs with directive verbs. Our initial predictions were largely borne out to be correct.

Table 3.4 summarises the key findings from this experimental study. In particular, we see that the only context in which British speakers allowed the subjunctive was in complement clauses to directive predicates. However, unlike desire and emotive-factive predicates where British English speakers demonstrated a categorical preference for one particular form over another (the *for* structure for desire verbs; indicative CPs for emotive-factive), the very same participants allowed both indicative- and subjunctive CPs as complements to directive verbs. This suggests that both types of CPs are acceptable under directive predicates.

It was unexpected, however, that British English speakers would only allow *for* infinitives with desire predicates, and not emotive-factive predicates, contra Emonds (1970, 1976) and Bresnan (1972). While further discussion of this is beyond the scope of the current thesis, it is possible that such a finding is a direct product of both diachronic changes and dialectal differences in the morphosyntactic expression of modality.

More generally, these findings indicate that L1 speakers of British English have access to the modal feature, [M], in the complement phrases of directive and desire (but not emotive-factive) predicates, albeit to varying degrees. The main difference between directive and desire predicates is how both verb types mark the interpretable modal feature, [M]: where directive predicates use inflectional morphology (more specifically, the zero morpheme), desire predicates use the preposition, *for*, to mark modality. As far as the current study shows, emotive-factive predicates do not appear to express modality morphosyntactically in any way.

Table 3.4 Distribution of forms-to-meanings in English

	Desire	Directive	Emotive-factive
For	✓	✗	✗
Indicative	✗	✓	✓
Subjunctive	✗	✓	✗

Chapter 4 Study 1: Obligatory Subjunctives

4.1 Introduction

Formal approaches to second language (L2) development have typically used offline measures to examine the role of the first language (L1) on L2 comprehension. Much less is known, however, about how L1 grammatical knowledge, particularly L1/L2 differences in feature configurations, influences L2 online processing. The subjunctive is a particularly interesting test case to investigate whether differences in L1/L2 featural composition modulate offline and online sensitivity to verbal mood. As will become clear in following sections, English and French exhibit crosslinguistic variation in terms of the syntactic and lexical-semantic properties governing verbal mood.

Previous research has consistently shown that learning how to express grammatical mood (indicative/subjunctive) is challenging for second language learners whose first language (e.g., English) does not grammaticalise the distinction to the same extent (Howard, 2008; Bartning, Lundell and Hancock, 2012; Ayoun, 2013; McManus and Mitchell, 2015). However, this research has typically used offline methods, such as judgment and production tasks, to probe L2 knowledge of the subjunctive. Few studies have examined this from a processing perspective, especially using online measures, such as self-paced reading, eye-tracking and ERPs.

The current chapter is organised as follow. It begins by discussing the distributional and interpretational properties of the French subjunctive in obligatory contexts (Section 4.2). In Section 4.3, we advance fine-grained, concrete predictions relating to the L2 acquisition of the French subjunctive by native speakers of British English within a processing-based extension of the Feature Reassembly framework, combining the findings from *Study 0* on the availability of the subjunctive in British English with our discussion on the French subjunctive from Section 4.2. The rationale for the current study is further explored in more detail through an examination of previous studies on the L2 acquisition of the subjunctive across Romance languages (Section 4.4). Section 4.5 presents the research questions as well as study-specific predictions, with information about the participants, the experimental design and methodology laid out in Section 4.6. The results of the study are described in Section 4.7. Finally, Section 4.8 discusses the implications of these results in light of the research questions, previous studies and current theoretical debates in the field.

4.2 Obligatory Subjunctive in French

The current study focuses on the use of the subjunctive in subordinate clauses where mood selection depends on the lexical-semantic properties of the matrix predicate. Following Portner (2018: 4), it is assumed that mood represents "an aspect of linguistic form which indicates how a proposition is used in the expression of a modal meaning". In this section, we describe subjunctive use in French and provide a selective overview of related theoretical analyses.

Broadly speaking, volitional subjunctives (henceforth, obligatory subjunctives) constitute the core class of subjunctive complements. Here, mood selection is triggered, in large part, by the lexical-semantic properties of the matrix predicate (Giannakidou, 1998, 2009; Quer, 2009). In French, verbs of saying (*dire* 'say', *observer* 'observe') and epistemic (*réaliser* 'realise', *se rappeler* 'remember') predicates typically select indicative complement phrases (CPs), whereas emotive-factive (e.g., *regretter* 'regret'), desire (e.g., *préférer* 'prefer', *souhaiter* 'wish', *vouloir* 'want') and directive (e.g., *suggérer* 'suggest', *insister* 'insist') predicates take subjunctive CPs (Baunaz, 2017). For example, in (20), the matrix verb, *penser* 'to think', licenses the indicative in the lower clause, whereas the matrix verb, *vouloir* 'to want', in (21), triggers the subjunctive in the lower clause.

(20) Les étudiants pensent que la première ministre est incompétente
The students think that the first minister is$_{IND}$ incompetent
'The students think that the prime minister is incompetent.'

(21) Le gouvernement veut que la 4G soit disponible partout
The government wants that the 4G is$_{SUBJ}$ available everywhere
'The government wants 4G to be available everywhere.'

Table 4.1 Distribution of veridicality,[13] mood and attitude in French

	Veridicality	**Mood**	**Attitude**
Emotive-Factive	Relative	Subjunctive	Emotive
Desire	Non-veridical	Subjunctive	Emotive
Directives	Non-veridical	Subjunctive	Emotive
Epistemic	Strong veridical	Indicative	Non-emotive
Verbs of Saying	Veridical	Indicative	Non-emotive

Adapted from (Baunaz, 2017, p. 22)

[13] "A propositional operator F is veridical iff from the truth of F_p we can infer that p is true according to some individual x (i.e., in some individual x's epistemic model." (Giannakidou, 1998, p. 1889)

The exact nature of the lexical-semantic properties is open to debate (Portner, 2018). A particularly interesting proposal is that of Baunaz (2017). In the spirit of Dowty (1991), Baunaz (2017) uses the principle of semantic compositionality to argue that predicates selecting tensed complementiser phrases (CPs) can be decomposed into four sub-events[14] (sentient, volitional, cause and emotive) depending on the semantic properties of the matrix predicate, as summarised in Table 4.2. These sub-events are, in a sense, comparable to semantic features. However, not all these features are shared by predicates selecting tensed CPs. For example, emotive-factive verbs only share the sentient feature with verbs of saying and epistemic verbs. In addition, the emotive feature is exclusively found in desire, directive and emotive-factive verbs; the volitional feature in desire and directive verbs but not emotive-factive verbs and the causal feature in directive verbs.

Table 4.2 Semantic decomposition of verbs selecting tensed CPs

	Saying	Epistemic	Emotive-factive	Desire	Directive
Mood	Ind.	Ind.	Subj.	Subj.	Subj.
Sentient	yes	yes	yes	yes	yes
Volitional	no	no	no	yes	yes
Cause	no	no	no	no	yes
Emotive	no	no	yes	yes	yes

Adapted from (Baunaz, 2017, p. 22)

It is important to note, however, that within the nano-syntactic framework (Caha, 2009; Starke, 2010, 2014), semantic features are hierarchically ranked in relation to the universal functional sequence of syntax (fseq). Consequently, directive verbs consist of four sub-events (i.e., casual, volitional, emotive and sentient). As we can see from the functional projection in (22), cause is the most relevant interpretation of these external arguments (that is, sub-events) for directive verbs.

(22) $[_{\text{CauseP}}$DP4 cause $[_{\text{VolitionalP}}$ DP3 vol $[_{\text{EmotiveP}}$ DP2 emo $[_{\text{SentienceP}}$ DP1 sent]]]]

(23) $[_{\text{VolitionalP}}$ DP3 vol $[_{\text{EmotiveP}}$ DP2 emo $[_{\text{SentienceP}}$ DP1 sent]]]

(24) $[_{\text{EmotiveP}}$ DP2 emo $[_{\text{SentienceP}}$ DP1 sent]]

[14] Ramchand (2008a, 2008b) suggests that (dynamic) predicates are broken downs into subparts of the universal functional sequence of syntax (fseq), in which "each projection corresponds to a subevent with its own predicational Subject position, and linked by the generalized 'leads-to' or 'cause' relation" (Ramchand, 2008a, p. 118). Each sub-event (semantic feature) of a verb is attached to a specific syntactic head.

(25) $[_{SentienceP}$ DP1 sent]

Desire verbs, on the other hand, consist of three stative sub-events (i.e., volitional, emotive and sentient), as illustrated in the functional projection in (23). Here, the subject of volition has to be emotive and sentient and express a volitional reaction towards the embedded CP, i.e., the object of volition. In contrast, emotive-factive verbs only involve two stative sub-events (i.e., emotive and sentient), as illustrated by the projection in (24). Here, the emotive feature is the most important semantic feature of the external arguments. Finally, verbs of saying and epistemic verbs consist of a single sub-event (i.e., sentient), as demonstrated in (25). Given that all predicates licensing subjunctive CPs share the emotive feature and those licensing indicative CPs do not, Baunaz (2017) proposes that it is the emotive feature that triggers mood selection in French.

An alternative proposal that attempts to account for mood choice in such contexts is the work of Quer and Giannakidou. Quer (1998, 2001, 2009), adopting Giannakidou's (2009) concepts of the individual model and anchoring, argues that all sentence-embedding predicates attribute an individual model relating to the predicate's accessibility relationship with their complement. However, unlike Giannakidou, who argues that mood selection is governed by veridicality, Quer claims that mood selection depends on whether there is a model shift. More specifically, the subjunctive is said to reflect a shift in the model evaluating the truth-value of the proposition from the default model, that is, the actual world as depicted in the speaker's epistemic model (Iverson, Kempchinsky and Rothman, 2008). In the complement clause of volitional predicates, this shift is triggered by the lexical-semantic properties of the selecting predicate. Here, the model of evaluation of volitional complement clauses is shifted from the default model to the bouletic model of the matrix subject. This consequently leads to "a shift both in the model's anchor (from the speaker to the matrix subject) and in the type of model (from an extensional epistemic model to a strongly intensional bouletic model, i.e., the set of worlds in which the desires of the matrix subject are realized)" (Iverson, Kempchinsky and Rothman, 2008, p. 138).

Based on this conceptual framework, Kempchinsky (2009) claims that the subjunctive is a purely syntactic phenomenon where subjunctive clauses consist of three functional projections: the Fin and Force heads in the left periphery and a Mood head in the inflectional field. The shift in the model of evaluation, as described in Quer's (2009) analysis, is expressed syntactically in the form of a modal feature, [M], in the Force head of the left periphery, which, in the case of obligatory subjunctives, is uninterpretable. Following Bianchi (2001), FinP represents the projection relating to the discourse context. In subjunctive clauses, FinP carries an uninterpretable modal feature inherited from Force; the deixis centre of the subjunctive is situated within the bouletic intensional model and "only indirectly linked to the extensional speech context" (Iverson, Kempchinsky and Rothman, 2008, p. 138). The interpretable feature of the MoodP head then values and deletes the uninterpretable modal features in the left periphery. Such an analysis is illustrated below:

(26) $\ldots\ V_{M}\ [_{CP}\ [_{ForceP}\ \text{Force}_{[uM]}\ [_{FinP}\ \text{Fin}_{[uM]}\ [_{IP}\ (\text{DP})\ [_{MoodP}\ [\text{V+T+M}][_{TP}\ldots]]]]]]$

selection   checking (agree)

(Iverson, Kempchinsky and Rothman, 2008, p. 139)

It is important to recognise that while the above analysis considers obligatory subjunctive complements to be a purely syntactic phenomenon, it equally recognises the implication of a semantic dimension. This semantic dimension stems from the lexical-semantic properties of the selecting predicate, whereas the presence of the morphosyntactic category of Mood in the complement clause represents the syntactic mechanism through which the predicate's selectional properties are expressed (Iverson, Kempchinsky and Rothman, 2008, p. 139).

Although the two analyses outlined above, namely those of Baunaz (2017) and Kempchinsky (2009), appear to reflect two disparate analyses of the subjunctive, the recognition that a semantic dimension is implicated allows us to unify the two. As Kempchinsky points out, the semantic dimension results from the lexical-semantic properties of the selecting predicate, and the lexical-semantic properties of the selecting predicate, in particular the question of semantic compositionality, is precisely what Baunaz (2017) describes. The current thesis therefore uses Baunaz's (2017) discussion of semantic compositionality to describe the semantic dimension of the subjunctive and Kempchinsky's analysis to reflect the syntactic dimension.

4.2.1 Distributional Properties

In the previous sub-section, we discussed how syntactic and lexical-semantic factors condition subjunctive use in French. This sub-section aims to explore the vitality of the subjunctive in French, focusing in particular on its distributional properties in the input.

It is widely reported that the French subjunctive, despite being an overt marker of mood, is infrequent in the input and lacks morphological transparency, particularly in comparison to the indicative (Poplack, 1989, 1990; O'Connor DiVito, 1997; Soutet, 2000; Poplack, Lealess and Dion, 2013). For example, although certain forms of the present subjunctive, particularly high frequency verbs, such as *être* 'to be', *avoir* 'to have' and *faire* 'to do', are morphologically distinct from the present indicative (e.g., *est*$_{\text{IND}}$ 'be' vs. *soit*$_{\text{SUBJ}}$ 'be'), other verbs do not exhibit observable morphosyntactic differences between the indicative and subjunctive forms (e.g., *chante*$_{\text{IND}}$ 'sing' vs. *chante*$_{\text{SUBJ}}$ 'sing'). According to Poplack (1989, 1990), the latter accounts for approximately one third of all subjunctive forms in Ontarian French.

In terms of frequency, Poplack (1989, 1990) reports that the subjunctive typically occurs between five and ten times for every thirty minutes of conversation. This research, however, has focused exclusively on the French spoken in Quebec, a variety widely assumed to differ markedly from metropolitan French (i.e., the French spoken in France). While this assumption has received little empirical attention, recent work by Kastronic (2016) sought to address the gap in the literature, while also exploring diachronic changes to the French language.

Kastronic (2016) found that in twentieth century Metropolitan French, the top-tier verbal governors[15] (i.e., *falloir, vouloir, aimer, penser (neg.)* and *croire (neg.))* accounted for approximately 53% of the governor pool, within which 92% of all embedded verbs were marked with subjunctive morphology. However, this rate dropped slightly in twenty-first century French, where these top-tier governors accounted for 47% of the governor pool, within which 87% of all embedded verbs were marked with subjunctive morphology. The study further observed that the lexical identity of embedded verbs played an important role with top-tier embedded verbs (i.e., *être*, *avoir*, *faire* and *aller*), representing 73% of the pool, of which 92% carried subjunctive morphology. Similarly, in twenty-first century French, we see that the top-tier embedded verbs accounted for 73% of the pool and 66% of all embedded verbs carried subjunctive morphology.

Table 4.3 presents the rate of subjunctive under the verbal governors included in Study 1. The vast majority of the verbal governors consistently select subjunctive CPs, with the exception of *croire (neg.)* and *penser (neg.)*, where we find an alternation between the indicative and the subjunctive. In other words, the subjunctive is more predictable than the indicative after certain verbal governors. In addition, it would appear that certain subjunctive triggers are more frequent in the input than others. The influence of these lexical properties, namely frequency and predictability, have important implications for the acquisition task, especially in light of our earlier discussion about how lexical properties can impact on processing patterns.

[15] A governor refers to a type of matrix predicate that licenses the subjunctive.

Table 4.3 Rate of subjunctive under verbal governors

		20th century			21st century		
Semantics	**Predicate**	**% subj**	**N subj**	**Total N**	**% subj**	**N subj**	**Total N**
epistemic	*croire* (neg.)	72	26	36	40	2	5
	penser (neg.)	72	36	50	58	14	24
desire	vouloir	96	54	56	90	27	30
	préférer	100	10	10	100	4	4
	souhaiter	100	2	2	100	3	3
directive	exiger	100	1	1	NA	NA	NA
	demander	100	3	3	100	1	1
	recommander	NA	NA	NA	NA	NA	NA
emotive-factive	aimer	95	18	19	79	11	14
	regretter	100	2	2	NA	NA	NA
	être + attribute	84	49	58	90	28	32
	craindre	100	2	2	NA	NA	NA

Note: NA represents a null observation.

4.3 Acquisition Task Predictions

Multiple theories have been advanced to account for the residual optionality (i.e., the variation and lack of convergence) that L2 learners demonstrate in their end-state grammars, including the Interpretability Hypothesis (Hawkins and Hattori, 2006; Tsimpli and Dimitrakopoulou, 2007), the Interface Hypothesis (Sorace and Filiaci, 2006; Sorace, 2011) and the Feature Reassembly Hypothesis (Lardiere, 2009). For the purpose of this PhD thesis, the acquisition task predictions presented in this section were formulated in line with a processing-based extension of the Feature Reassembly Hypothesis, the analysis of the subjunctive presented in Section 4.2 and the findings from Study 0 on the availability of the subjunctive in British English.

To recap, the Feature Reassembly Hypothesis (Lardiere, 2009) stipulates that L2 learners begin language learning with "an entrenched system of morphosyntactic features already assembled into lexical items" (Hwang and Lardiere, 2013, p. 58). In order to acquire L2-specific features successfully, learners must undergo two stages. First, they must perform a contrastive analysis between L1 and L2 properties in an attempt to map, either accurately or inaccurately, the L1 form(s) onto the L2 equivalent(s). If L1/L2 form-to-meaning mapping mismatches occur, L2 learners must reassemble existing L1 feature configurations into new L2 feature bundles and lexical items (Hwang and Lardiere, 2013, p. 58). Linguistic properties with divergent L1/L2 feature configurations are expected to be problematic given the need for feature reassembly. This hypothesis has found empirical support from recent studies by Guijarro-

Fuentes (2012), Gil and Marsden (2013), Hwang and Lardiere (2013), Cho and Slabakova (2014), Domínguez, Arche and Myles (2017), among others.

As we established in Study 0, British English speakers license the subjunctive in their L1 grammar, but only in complements to directive verbs and in alternation with indicative CPs. This means that these learners should already have the functional category, MoodP, and all the relevant modal features in place. Their task is to map the L1 forms to the corresponding lexical items in L2 (i.e., subjunctive morphology) by acquiring new (overt) morphology via the head of MoodP. English learners of French will then have to learn that the distribution between morphosyntactic form and syntactic contexts differs in the two languages and thus widen the syntactic contexts for the projection of TP to subjunctive clauses in French.

For this to happen, feature remapping is required. More specifically, in order to acquire the appropriate syntactic licensing contexts for the subjunctive, learners of French are required to redistribute the semantic features associated with verbs selecting indicative CPs or *for-to* infinitives onto verbs selecting subjunctive CPs. The specific mapping of semantic features onto lexical items in English and French is summarised visually in Figure 4.1.

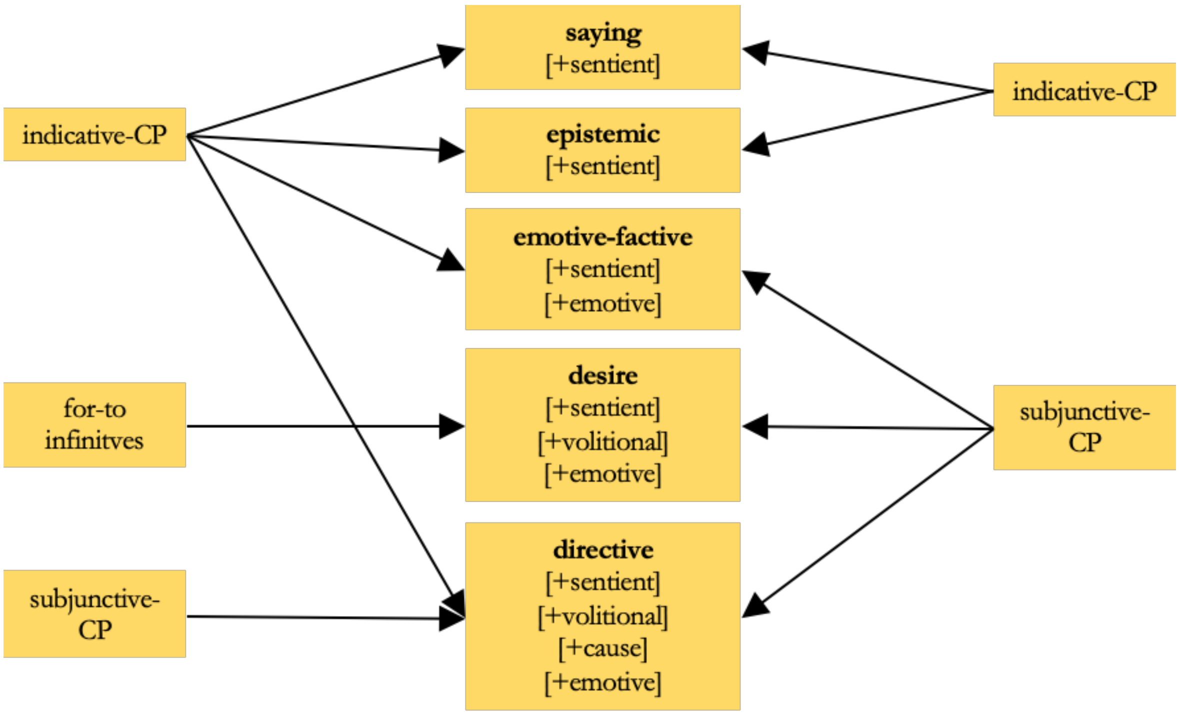


Figure 4.1 Mapping of the semantic features of verbs selecting tensed CPs (middle column) onto corresponding forms in English (left column) and French (right column)

According to this analysis, English learners of French need to (1) learn that each morphological form available in French (indicative and subjunctive) is mapped onto a specific lexical-semantic category and its feature bundle without ambiguity and (2) redistribute the semantic feature bundles associated with indicative CPs, subjunctive CPs and *for-to* infinitives in English onto these two forms.

More specifically, when acquiring French mood morphology, English speakers must learn that the distribution of form-to-meaning mappings differ in French and English. In particular, that indicative CPs can only be mapped onto verbs of saying and epistemic verbs (and their feature bundles), but not emotive-factive and directive verbs in French, as would be the case in English. Instead, learners are required to remap the semantic feature bundles associated with emotive-factive verbs onto subjunctive CPs from indicative CPs. With directive verbs, L2 learners must move from a superset relation in the L1, where both indicative and subjunctive CPs are allowed as complements to directive verbs, to a subset relation in the L2, where only subjunctive CPs are permitted. For desire verbs, learners must remap the semantic feature bundle associated with this verb class onto subjunctive CPs from *for* infinitives.

In accordance with Full Transfer, we predict that English speakers acquiring French will initially allow both indicative and subjunctive CPs as complements to directive predicates since they are the closest morpholexical equivalent of the two forms. For other verbs, we anticipate that these learners will initially assume that French indicative CPs are the morpholexical equivalent of English tensed CPs. For verbs selecting subjunctive CPs, therefore, at initial or early stages of L2 development, we hypothesise that English native speakers will alternate between indicative and subjunctive CPs for directive verbs, due to L1 influence, but overuse indicative CPs for all other verbs selecting tensed CPs in French. Negative evidence is then necessary to unlearn that these other verbs do not select indicative CPs in the L2.

At intermediate to advanced proficiency levels, we predict that learners will exhibit more target-like sensitivity to subjunctive complements with directive verbs first, and then will gradually remap the semantic feature bundles of the lexical-semantic properties of desire and emotive-factive verbs to allow for the selection of subjunctive CPs. It is possible that more target-like sensitivity will surface with desire predicates first in comparison to emotive-factive verbs. As we have already discussed in Section 3.2.1, *for-to* infinitives have been analysed as markers of modality. This means that L2 learners already have access to the modal feature, [M], via the preposition, *for*, but must instead remap this feature onto inflectional morphology via MoodP. This is not the case with emotive-factive verbs which only accept indicative CPs in the L1.

It must be recognised, however, that the success of this reassembly process is ultimately contingent on the abundance of unambiguous input, which in the case of the French subjunctive is not certain. Slabakova (2015, p. 680) argues that the input plays a crucial role in "setting parameter values within the generative framework" and that successful acquisition depends, in large part, on the quality and quantity of exposure. Constructions that are not salient and/or frequent in the input are therefore expected to be more problematic for L2 learners. The input is likely to be a complicating factor in the case of the French subjunctive given its lack of salience and infrequency (O'Connor DiVito, 1997; Poplack, Lealess and Dion, 2013).

4.4 Studies on L2 Acquisition of Obligatory Subjunctive

As we discussed briefly in the introduction to this thesis, many studies have shown that the French subjunctive is a source of difficulty for learners whose first language (e.g., English) does not fully grammaticalise this distinction (Iverson, Kempchinsky and Rothman, 2008; Ayoun, 2013; Borgonovo, Garavito and Prévost, 2015; McManus and Mitchell, 2015). In this section, we review these studies in more depth and explore the factors contributing to these difficulties. In the final part, we discuss why eye-tracking represents a valuable tool to explore the acquisition of the French subjunctive.

Most studies to date have used primarily offline measures of grammatical knowledge to assess learners' production and comprehension of the French subjunctive. Data elicitation techniques include the production of personal narratives, sentence completion, (semi-structured) interviews and interpretation tasks, just to name a few. Ayoun (2013), for example, collected cross-sectional data from 42 university learners across three proficiency levels: beginners (n = 14), intermediate (n = 15) and advanced (n = 13). In order to assess use and development of the French subjunctive, participants completed personal narratives and a sentence completion task. Results from the personal narratives demonstrated infrequent but accurate use of the subjunctive across all three proficiency levels; a finding consistent with previous studies involving oral tasks (Bartning, 2008; Howard, 2008). Bartning suggested that L2 learners' infrequent use may relate, at least in part, to the infrequency and ambiguity of subjunctive forms in the native speaker input rather than non-target-like behaviour. Ayoun, however, observed distinct proficiency effects in the sentence completion task with respect to accuracy: 19.9% (beginners), 50.3% (intermediate) and 60.9% (advanced). Given the discrepancy in findings between the two tasks (personal narratives vs. sentence completion), Ayoun argued that task type may have influenced learners' subjunctive knowledge, in that learners had more freedom over which forms to produce in the personal narratives. This finding suggests that task design can significantly influence the performance of learners and any resulting conclusions.

The deterministic role of proficiency has also been documented in various studies, with more accurate performance reported among advanced learners than any other proficiency group (Ayoun, 2013; McManus and Mitchell, 2015), particularly in terms of the range of forms and types of subjunctive triggers employed. For example, Ayoun found evidence of subjunctive use with conjunctions and affirmative clauses across all L2 learner groups, but that the variety of triggers used increased with proficiency. These findings echo those of Bartning (2008) and Howard (2008) who both observed that learners used the subjunctive in affirmative clauses, in particular *falloir que* and *vouloir que,* more frequently than any other trigger.

The lexical-semantic properties of subjunctive triggers, and by extension the L1, also appear to modulate accurate subjunctive use in affirmative clauses to a certain extent. For example, Ayoun (2013), in a sentence completion task, tested L2 learners of different proficiencies on their knowledge of subjunctive triggers, grouped according to the lexical-semantic properties of the matrix verb. These categories included: judgment, emotion, wish/regret, order/interdiction and doubt/impossibility. It is particularly interesting that all L2 learners, regardless of proficiency, performed markedly better with order/interdiction triggers in comparison to any other trigger. Such a finding suggests a potential role for the L1. Recall that in Section 4.3, we predicted that L2 learners would be more likely to exhibit sensitivity effects with directive predicates than any other lexical-semantic category since the L1 grammar also licenses the use of the subjunctive in such contexts. The results from Ayoun's study are therefore consistent with these predictions.

The role of learning context, more specifically an extended residence in the target language community, has also been investigated. In a longitudinal design, McManus and colleagues (Mcmanus *et al.*, 2014; McManus and Mitchell, 2015), examined how subjunctive use develops over the course of a residence abroad experience, by analysing data from two production tasks (oral and written) and a grammaticality judgment task. The authors observed that during the extended residence abroad, learners, particularly those of lower proficiencies, developed more markedly in non-variable contexts (e.g., in affirmative clauses) in comparison to variable contexts (e.g., with epistemics in negated clauses), but that ultimately residence abroad had a limited effect on L2 subjunctive use.

Howard (2008) presented similar findings using sociolinguistic interviews with English-speaking university learners (n = 18). Results showed that formal instruction improved subjunctive use, but that this development was restricted to a limited number of triggers and forms. Study abroad, however, appeared to widen learners' range of forms, potentially due to increased exposure to the naturalistic input. These findings suggest that learning context plays a significant role in the development of subjunctive use in specific domains.

Although the current section has largely focused on the acquisition of the French subjunctive, a significant body of research has also explored the acquisition of the Spanish subjunctive. Notable studies include Borgonovo and Prévost (2003), Iverson *et al.* (2008) and Borgonovo *et al.* (2015). While these studies examined different uses of the subjunctive, we only focus on the results pertaining to the obligatory subjunctive.

In fact, the study by Iverson and colleagues is one of the few studies to investigate knowledge of the subjunctive[16] within a formal linguistics framework among intermediate and advanced L2 learners of Spanish. To briefly summarise their analysis of the obligatory subjunctive in English and Spanish, the authors propose that English learners of L2 Spanish must retrieve the functional projection, MoodP, and an uninterpretable modal feature on the head of Force via the formal English register of the L1 grammar, whilst also unacquiring (delearning) the structure of the English *for-to* clause. Such an analysis therefore suggests that successful acquisition crucially depends on the availability of the MoodP projection in the L1 grammar (for a more in-depth discussion of this analysis applied to the L2 acquisition of the French subjunctive, see Section 4.3). Based on such an analysis and findings from a morphology recognition task and a scalar grammaticality judgment task, Iverson *et al.* (2008, p. 156) argued that the advanced learners had clearly acquired obligatory subjunctive complements, with the intermediate learners demonstrating "strong indications of being successful". In short, the study revealed that obligatory subjunctives are not necessarily a source of persistent difficulty for English-speaking L2 learners of Spanish.

Only very recently, however, have researchers started to use online methods, such as ERP, self-paced reading or eye-tracking, to study the real-time processing of the subjunctive (Cameron, 2011; Villegas Erce, 2014). For example, Cameron's (2011) PhD thesis investigated whether non-native speakers of Spanish can process inflectional mood morphology in a native-like fashion during real-time sentence comprehension. Native speakers of Spanish and L2 learners across three proficiency levels (intermediate, upper intermediate and advanced) completed a self-paced reading task. Results showed that native speakers demonstrated

[16] Note that the main focus of this study was on knowledge of the indicative/subjunctive distinction with negated epistemic predicates. The results of this will be discussed in Section 5.2.2.

sensitivity (i.e., increased reading times) to modality-mood mismatches, but L2 learners did not, regardless of proficiency level.

Although not directly relevant to the current study, it is interesting to report findings from Villegas Erce's (2014) thesis, given the paucity of research using eye-tracking as a methodology. Villegas Erce examined the role of L2 English immersion on the processing of L1 Spanish sentence complement/relative clause ambiguities, using the Spanish mood system (the indicative/subjunctive) as a tool to study changes to the L1. Results from an eye-tracking during reading experiment showed no significant differences in processing patterns between three groups of native speakers of Spanish in different L2 immersion conditions (non-immersed, immersed and heritage speakers) and that these speakers were able to use the information encoded in the verb to anticipate verbal mood morphology in the embedded clause.

The use of eye movement data, especially when combined with offline data, can provide L2 researchers with a rich insight into the development of the French subjunctive, which offline data alone cannot. A commonly cited criticism of traditionally used offline tasks, such as acceptability judgments, is that they rely predominately on automatized, explicit knowledge (Suzuki, 2017). This is particularly problematic when studying the French subjunctive, a structure about which (instructed) L2 learners possess high levels of explicitly learned, metalinguistic knowledge. Online methods, on the other hand, allow us to measure interpretation in real-time and thus tap into L2 learners' underlying implicit knowledge of language (Keating and Jegerski, 2015). In theory, eye movement data should reduce learners' reliance on explicitly learned knowledge, and limit the amount of time available for conscious linguistic problem solving (Keating and Jegerski, 2015).

Due to the high spatial-temporal resolution information it offers, eye movement data can help us to understand the exact location that any disruptions to language comprehension occurs and the subsequent impact it has on the time course of language processing. For example, when processing the subjunctive, speakers must integrate information from the matrix predicate (i.e., lexical-semantic information) and the wider context (i.e., discourse-pragmatic information) in order to accurately anticipate verbal mood morphology in the embedded clause. Eye movement data allows us to measure the extent to which L2 learners are successful in achieving this by analysing a combination of early and late measures. For example, early measures, such as gaze duration, can tell us whether readers are sensitive to syntactic and semantic anomalies (i.e., mood–modality incongruencies), whereas late measures, such as rereading of a region of interest (in this case, the embedded verb), can give an indication of time spent resolving or reanalysing any earlier integration difficulties (Rayner *et al.*, 2004).

4.5 Research Questions and Predictions

The current study aims to investigate the influence of the L1 morphosyntax on the processing of mood–modality mismatches in L2 French at different levels of proficiency and residence in a French speaking environment. With this in mind, we advance the following research questions and related predictions:

- Do English-speaking L2 learners of French exhibit offline and online sensitivity to the syntactic and semantic constraints of the subjunctive in obligatory contexts?

- To what extent is this sensitivity influenced by L1/L2 differences in
    - feature configurations,
    - proficiency and
    - residence in a French-speaking country?

We expect to observe significant differences between indicative and subjunctive sentences, among the L1 speakers, in each of the reading time measures at both the critical and spillover region, with longer reading times for the ungrammatical indicative than the grammatical subjunctive sentences, regardless of the matrix predicate's lexical-semantic properties. Our prediction regarding longer reading times for ungrammatical (than grammatical) sentences is based on previous research showing that certain reading times, most notably gaze duration, are sensitive to morphosyntactic violations (Rayner *et al.*, 2004).

In keeping with the theories and research presented in the previous sections, we anticipate that the L2 group will demonstrate an asymmetry between offline and online measures with increased sensitivity to mood–modality mismatches in the offline than the online data (cf., Papadopoulou and Clahsen, 2003; Marinis *et al.*, 2005; Shimanskaya and Slabakova, 2017; Stepanov *et al.*, 2019). Furthermore, we expect to find observable differences in processing patterns between the native speaker control group and the L2 group that are modulated by L1/L2 morphosyntactic differences. More specifically, we should see more target-like sensitivity to mood–modality mismatches with directive predicates than with desire and emotive-factive predicates due to the feature remapping and reassembly differences between tensed (indicative/subjunctive)-CP-selecting predicates in the L1 and L2, as discussed in Section 4.3. Under the Feature Reassembly Hypothesis, proficiency and residence abroad are predicted to modulate L2 processing patterns. In particular, as learners become more proficient and are exposed to more naturalistic input through residence abroad, sensitivity should become more target-like.

Among the less proficient L2 speakers, particularly the non-resident abroad learners, we should not find any significant differences between indicative and subjunctive sentences with directive verbs, regardless of the reading time measure or region (critical vs. spillover). With emotive-factive and desire verbs, we are likely to detect a significant difference between indicative and subjunctive sentences on each of the reading time measures at both the critical and spillover region, with longer reading times for subjunctive than indicative sentences, contra our prediction for L1 speakers. As proficiency increases, we hypothesise that learners will exhibit significantly longer reading times for indicative than subjunctive sentences, initially with directive verbs and then later with emotive-factive and desire verbs. If L2 speakers do demonstrate target-like sensitivity to mood–modality mismatches, we expect this effect to be more evident in the spillover than the critical region. This prediction is based on previous research that has shown that sensitivity to morphosyntactic inconsistencies is often delayed among L2 learners (Felser and Cunnings, 2012; Felser *et al.*, 2012; Lim and Christianson, 2015; Boxell and Felser, 2017).

(27) Summary of predictions:

- Hypothesis 1: L2 learners should be more sensitive to mood–modality mismatches in the offline data than the online data.
- Hypothesis 2: If sensitivity to mood–modality mismatches is modulated by the L1 properties, then L2 learners should demonstrate sensitivity first with directive predicates, then desire predicates and finally emotive-factive predicates.
- Hypothesis 3: As proficiency increases, L2 learners should exhibit sensitivity to mood–modality mismatches, with not only directive predicates, but also desire and emotive-factive predicates.
- Hypothesis 4: Exposure to (naturalistic) input, through residence abroad, should improve sensitivity to mood–modality mismatches.

4.6 Method

4.6.1 Participants

We tested a total of 75 participants, including a control group of 30 L1 speakers of French and a test group of 45 (English-speaking) L2 learners of French. All participants reported normal or corrected-to-normal vision and were recruited from the student population at the University of Southampton. They were paid £10.00 for their participation in the study. Any individual who reported more than one L1 or whose L1 was not English or French were excluded from the analysis. In light of previous research into the effects of educational attainment and other socioeconomic factors on language outcomes (Pakulak and Neville, 2010; Mulder and Hulstijn, 2011), we restricted recruitment to participants from similar educational and socioeconomic backgrounds to the L2 group, i.e., students or recent graduates with at least two years of university education.

Language History Questionnaire

Table 4.4 Background information for participants. Means (standard deviations)

		L1 speakers (n = 30)	L2 speakers (n = 45)
Proficiency (LexTALE) Score		86.89 (6.78)	60.82 (7.40)
Age		21.82 (1.76)	21.42 (1.12)
Gender	Male	8	9
	Female	24	34
Education	Undergraduate	23	40
	Postgraduate (Masters)	7	3
	Postgraduate (PhD)	2	0
Handedness	Left	3	5
	Right	29	38
Current City	Southampton	30	45
Residence in a French-speaking country		30	20
Months spent in a French-speaking country		NA	12.95 (11.50)
Months in Southampton, UK		7.73 (10.57)	NA
Years of English Use		11.5 (4)	20.97 (1.50)
Years of French Use		21.04 (2.87)	11.08 (3.33)
Self-assessed Proficiency in English		0.73 (0.07)	NA
Self-assessed Proficiency in French		NA	0.78 (0.08)
English Immersion Quotient		0.28 (0.06)	0.46 (0.01)
French Immersion Quotient		0.44 (0.02)	0.27 (0.04)
English Dominance Quotient		0.45 (0.07)	NA
French Dominance Quotient		0.52 (0.11)	0.47 (0.11)

Prior to testing, participants were asked to complete a modified version of the Language History Questionnaire (LHQ 3.0) in either English or French, depending on their preferred language (Li *et al.*, 2019). This questionnaire allowed us to ascertain linguistic and demographic information about the participants. Table 4.4 and Table 4.5 summarise the data.

Based on this data, we were able to calculate immersion and dominance quotients. The immersion quotient for French, and where possible English, was calculated using the participant's age, age of acquisition and years of language use, whereas the dominance quotient was based on participant's self-reported proficiency and the time in hours that the participant spent on different aspects of each language. In addition to the LHQ, L2 participants who had completed an extended residence abroad less than a year before testing were asked to complete the Language Exposure Questionnaire (LEQ) in order to retrospectively assess their language use during their residence abroad (Mitchell, Tracy-Ventura and McManus, 2017).

Table 4.5 Participants' age of acquisition and self-assessed proficiency.

			L1 speakers (*n* = 30)	L2 speakers (*n* = 45)
Age of Acquisition	Reading	English	9.93 (2.67)	2.95 (1.39)
		French	5.46 (1.56)	10.03 (1.80)
	Writing	English	8.67 (5.45)	9.60 (3.91)
		French	5.45 (2.02)	10.28 (1.76)
	Speaking	English	9.30 (2.97)	0.00 (0.00)
		French	0.09 (0.53)	9.47 (2.41)
	Listening	English	9.30 (3.01)	0.00 (0.00)
		French	0.00 (0.00)	9.69 (2.38)
Self-Assessed Proficiency	Reading	English	5.29 (0.66)	6.00 (0.00)
		French	6.00 (0.00)	5.63 (0.59)
	Writing	English	5.18 (0.55)	6.00 (0.00)
		French	6.00 (0.00)	5.30 (0.76)
	Speaking	English	4.96 (0.58)	6.00 (0.00)
		French	6.00 (0.00)	5.40 (0.87)
	Listening	English	5.11 (0.63)	6.00 (0.00)
		French	6.00 (0.00)	5.55 (0.60)

Acquisition and processing studies have historically sought to recruit almost exclusively monolingual speakers as a control group. However, the vast majority of the L1 control group in this study had been resident in the United Kingdom for an average of 7.73 months at the time of testing. It was our intention to recruit L1 French speakers whose L2 competence in English was comparable to that of our L2 learners in French. Research has shown that the parallel activation of two languages in the mind leads to competition that places cognitive demands on the speaker to suppress the language not in use in order to process the target language (e.g., Green and Abutalebi, 2013; Kroll and Bialystok, 2013; Baum and Titone, 2014). The cross-language competition is argued to occur at every level of language processing and has been detected in reading, listening and speaking (e.g., Marian and Spivey, 2003; Dijkstra, 2005; Kroll, Bobb and Wodniecka, 2006).

Such research has implications for the debate on the comparative fallacy of comparing L2 learners with monolingual L1 speakers (Bley-Vroman, 1983). As Schmid, Gilbers and Nota (2014, p. 152) state, when you compare learners with monolingual native speakers, you are "not only asking mere mortals [i.e., L2ers] to run as fast as Usain Bolt [i.e., monolingual natives], [you are] asking them to do so with lead weights [i.e., the L1] attached to their feet." With this in mind, we recruited bilingual speakers as our test group and conducted separate statistical analyses for L1 and L2 speakers.

4.6.2 Independent Proficiency Test

In order to assess participants' overall proficiency levels in French, we administered the French version of LexTALE (Brysbaert, 2013), based on the original English version by Lemhöfer and Broersma (2012). LexTALE is a standardised test of vocabulary knowledge aimed at learners of upper-intermediate to advanced proficiency. Results from this test have been found to strongly correlate with standardised measures of proficiency, such as the Quick Placement Test (QPT), and thus provide an accurate indication of overall proficiency.

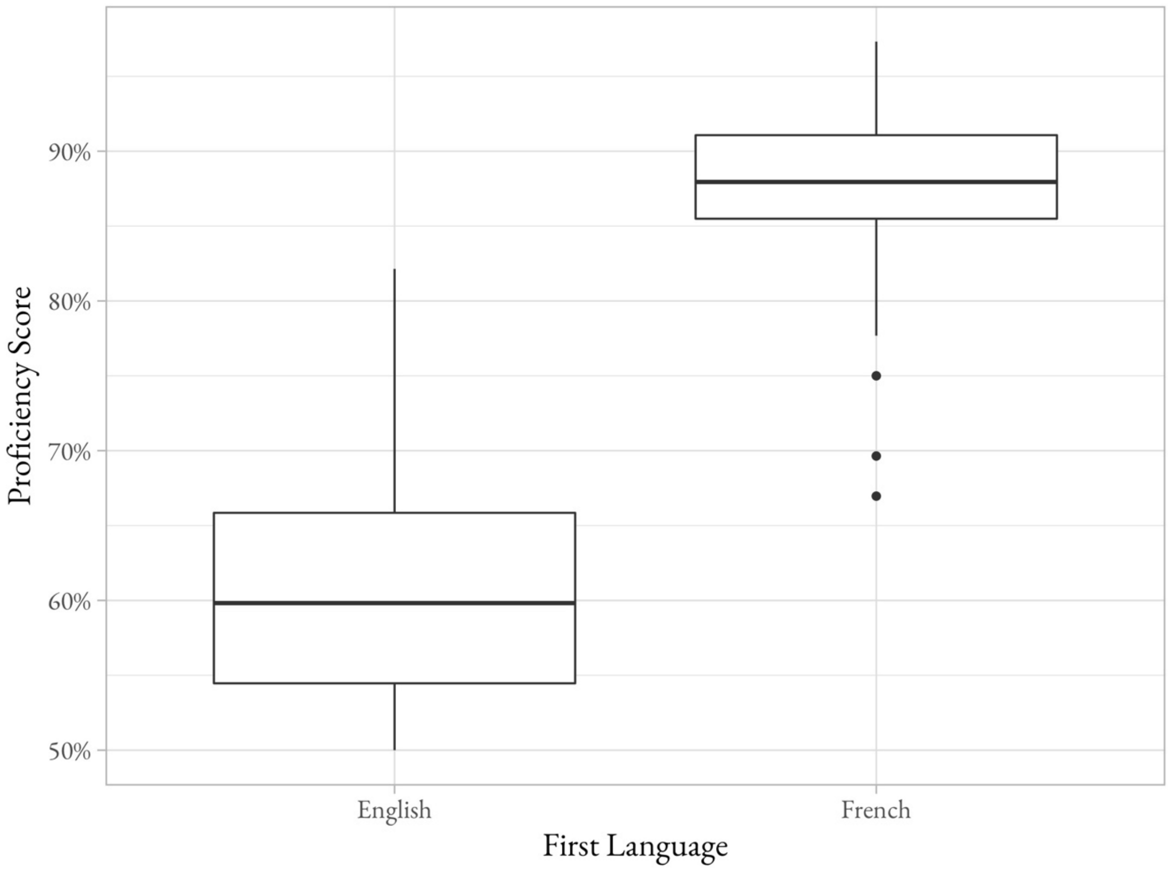


Figure 4.2 Proficiency scores from LexTALE test

The task in question consisted of an untimed lexical decision task in which participants were presented with a list of 56 French words of varying frequencies and 28 French-looking nonwords. Participants were required to identify which of the 84 words were French. On average, this task was completed in 3 minutes. Figure 4.2 provides a visual overview of the L1 and L2 performance on the LexTALE test. A one-way ANOVA revealed a significant difference between the two groups ($F(1,76) = 257.1, p > .001$).

4.6.3 Morphology Recognition Task

All participants completed a mood morphology recognition task. This task was designed to ensure that participants were able to differentiate between the indicative and subjunctive forms of the verbs used in the main tasks. The 22-item test was a shortened version of the sentence completion task used in Ayoun (2013), in which participants were asked to supply the appropriate subjunctive form of the lexical verbs provided in the bracket. The matrix predicates included in this task only required subjunctive complements and spanned five (broad) semantic categories, including emotion, judgment, order/interdiction, wish/regret and uncertainty/ doubt/(im)possibility/(im)probability. The embedded verbs consisted of both regular and irregular forms.

Table 4.6 Raw scores on morphology production task

	Morphology Task Score (%)
L1 speakers	95.45 (6.84)
L2 speakers	57.91 (26.66)

The average score on the morphology recognition task for each group is presented in Table 4.6. It shows that L1 speakers of French were able to consistently and accurately supply the appropriate subjunctive form. However, this was not always the case for L2 speakers. Results from the mixed-effects linear regression model presented in Table 4.7 revealed that proficiency, but not an extended residence in a French-speaking country, modulated their performance on this task. In particular, we found that as proficiency increased, L2 performance gradually became more target-like.

Table 4.7 Mixed-effects model on morphology production data.

	b	*SE*	*t*
(Intercept)	0.64	0.06	10.42
Proficiency	0.20	0.08	2.55
Residence	-0.07	0.08	-0.94
Proficiency x Residence	-0.13	0.09	-1.47

Note: Shaded cells denote statistical significance ($p < 0.05$).

These findings are surprising. Based on previous research suggesting that advanced L2 learners generally do not experience significant difficulties with the subjunctive in obligatory contexts (Ayoun, 2013; McManus and Mitchell, 2015), we predicted that participants would have been familiar with the subjunctive triggers included in this study and consistently supplied the appropriate subjunctive form. Initially, we planned to exclude participants who scored below 80% from the dataset. However, this would have resulted in significant data loss.

4.6.4 Eye Tracking Experiment Task

In order to examine L2 processing patterns, we conducted two eye-tracking during reading experiments. The first experiment (Study 1a) required participants to read a sentence and, after half of all trials, answer a comprehension question. The stimuli in the second experiment (Study 1b) was identical to the stimuli in Study 1a; the only difference was that participants were asked to rate its acceptability of each sentence. Both experiments are described in full in the next section.

4.6.4.1 Materials and Design

Table 4.8 Example of test items in obligatory contexts

Semantic Properties	**Mood**	**Example**
Desire	Indicative	*Cécilia veut que son amie **a** plus de confiance en elle. 'Cécilia wants her friend to have more faith in her'.
Desire	Subjunctive	Cécilia veut que son amie **ait** plus de confiance en elle. 'Cécilia wants her friend to have more faith in her.'
Directive	Indicative	*Amnesty demande que la police **fait** preuve de modération lors des manifestations. 'Amnesty requests that the police exercise restraint during the protests.'
Directive	Subjunctive	Amnesty demande que la police **fasse** preuve de modération lors des manifestations. 'Amnesty requests that the police exercise restraint during the protests.'
Emotive-Factive	Indicative	* Pierre craint que Sophie **est** toujours amoureuse de Claude. 'Pierre is afraid that Sophie is still in love with Claude.'
Emotive-Factive	Subjunctive	Pierre craint que Sophie **soit** toujours amoureuse de Claude. 'Pierre is afraid that Sophie is still in love with Claude.'

In Study 1, we recorded participants' eye movements while they read sentences, as in Table 4.8. All test items were bi-clausal sentences and followed the same underlying syntactic structure up until the critical region (i.e., the verb in the embedded clause), consisting of a singular third-person [+ animate] noun, a subjunctive CP licensing predicate and the complementiser *que* 'that' in the matrix clause followed by a singular third-person [+ animate] noun, the critical verb (i.e., *être* 'to be', *faire* 'to do', or *avoir* 'to have') marked with the condition-appropriate mood morphology (indicative or subjunctive) and the subsequent continuation in the embedded clause.

For each test item, we manipulated two within-subject factors. The first was the lexical-semantic properties of the matrix predicate, which had three levels: desire, emotive-factive and directive. The second factor was verbal mood (i.e., indicative and subjunctive) in the embedded clause. In each case, only the subjunctive form was grammatical. In total, we created 72 experimental items, such that participants saw 6 items per condition.

4.6.4.2 Lexical Properties of Test Items

Table 7.1 in Appendix A.2.1 summarises the lexical characteristics of the critical regions in this study. Recall that in Section 2.2.1, we observed that frequently, length and predictability can influence reading times. For this reason, eye-tracking researchers typically attempt to match critical regions across conditions and items on these lexical properties.

The ANOVA results in Table 7.2 in Appendix A.2.1 revealed significant differences between conditions for critical region length and zipf frequency (film subtitles and books). We argue that these differences were, to a certain extent, unavoidable due to the fact that the French subjunctive is an infrequent linguistic structure in the input whose morphosyntactic forms are typically longer in characters than indicative forms. For this reason, and to mitigate against any potential confounds, we included critical region length and/or frequency as control variables in our statistical models.

4.6.4.3 Piloting

To ensure that the test items exhibited the expected range of acceptability and rule out any unexpected differences in the material, we conducted an acceptability judgment pre-test via Amazon Mechanical Turk, a web-based survey tool. We must, however, interpret these judgment scores with caution since it is not possible to guarantee with absolute certainty that all participants were native speakers given the nature of the platform. This test was completed by 10 self-identified native speakers of French, none of whom participated in the main study. Participants were required to evaluate the grammaticality of each sentences on a scale of 1 to 7.

Table 4.9 Mean ratings for AMT pre-test (standard deviations)

	Indicative	Subjunctive
Desire	3.92 (2.22)	6.75 (0.55)
Directive	4.05 (2.04)	6.60 (0.90)
Emotive-factive	4.22 (2.18)	6.36 (1.29)

The descriptive statistics and estimates from the mixed-effects linear regression model, presented in Table 4.9 and Table 4.10, respectively, revealed a significant main effect for condition, but not for matrix semantic property. We further observed a significant interaction between condition and matrix semantic property (emotive-factive vs. directive). These results suggest that native speakers reliably distinguished between grammatical (subjunctive) and ungrammatical (indicative) sentences, regardless of matrix semantic property. We can thus conclude that the test items exhibited the expected range of acceptability.

Table 4.10 Mixed-effects model on pilot test scores

	b	*SE*	*t*
(Intercept)	5.29	0.21	25.33
Mood: Subjunctive vs. Indicative	1.26	0.07	17.37
Semantic: Desire vs. Directive	0.09	0.16	0.57
Semantic: Emotive-factive vs. Directive	0.01	0.16	0.05
Mood × Semantic (Desire vs. Directive)	0.17	0.09	1.77
Mood × Semantic (Emotive-Factive vs. Directive)	-0.20	0.10	-2.04

Note: Shaded cells denote statistical significance.

4.6.4.4 Design

From an initial set of 60, 36 test items were selected[17] following piloting, as described in the previous section. Each test item consisted of two versions (indicative CP and subjunctive CP). This resulted in a total of 72 test sentences. These items were then pseudo-randomised and distributed across two presentation lists in a Latin-square design and interspersed with an equal number of filler items, such that each participant saw a total of 72 sentences. The main experiment was divided into two blocks to allow participants to take a break if needed. Six practice items were included before the main experiment to familiarise participants with the procedure.

[17] These 36 test items were chosen on the basis of the grammaticality of the subjunctive (i.e., the grammatical) sentences, as tested in the pre-test. In other words, only the 36 most grammatical items were selected for inclusion.

4.6.4.5 Apparatus

Eye movements were recorded using the SR Research EyeLink 1000 Plus system with a sampling rate of 1000 Hz as participants read the stimuli on a 24-inch CRT monitor with a refresh rate of 60 Hz. The sentences were displayed in black monospaced font of approximately 22 pt on a white background. Although viewing was binocular, eye movements were recorded from the right eye. Participants were seated approximately 90 cm from the monitor and head movements were minimized using a chin and forehead rest.

4.6.4.6 Procedure

Both experiments were identical with respect to stimuli. The only difference between the two studies was the type of secondary task used. The order in which the two parts of the experiments took place was fixed and thus not counterbalanced, with the judgment task always following the comprehension task.[18]

In Study 1a, participants were instructed to silently read each sentence for comprehension at their normal reading rate (for no more than 10 seconds) and then answer a comprehension question (within 15 seconds) on 50% of the trials. Comprehension questions focused on the content of the sentence, not the form, and required a yes-no push-button response.

In Study 1b, participants were asked to silently read each sentence for a maximum of 20 seconds and then judge how natural the sentence sounded by selecting a number between 1 and 7, where '1' indicated that the sentence was completely unacceptable and '7' completely acceptable. An acceptability judgment task was included to probe speakers' offline knowledge of the subjunctive.

Both experiments started with a nine-point calibration procedure followed by six practice trials and a total of 36 experimental trials, which were presented in an individually randomized order and intermixed with 36 filler items. Participants saw a total of 72 sentences. At the start of each trial, a fixation point appeared at the location of the first letter of the sentence. Once

---

[18] We acknowledge that this will have led to repetition effects and quite possibly shorted overall reading times in the judgment task due to repeated exposure to experimental stimuli. However, we consciously adopted this order since the comprehension task was designed to reflect more closely natural reading, allowing participants to read for comprehension rather than for form. In this sense, it was designed to be the more implicit of the two tasks. Had the judgment task preceded the comprehension task, it is likely that the participants would have focused more on grammatical form in the comprehension task, which would have contradicted the overall goal of the comprehension task.

the participant's fixation on this point was stable, the sentence was displayed. Where necessary, recalibration was performed to compensate for any drift in calibration. We aimed for a mean calibration of 0.5 degrees of visual angle. The entire experiment lasted approximately 10 minutes.

4.6.4.7 Regions of Interest

Eye movements were analysed at the critical (the embedded verb in all conditions) and spillover region (one to three words of between 8 and 17 characters in total).[19] The decision to analyse reading times at the spillover region was motivated by previous research showing that sensitivity effects are often delayed, particularly among non-native populations (Felser and Cunnings, 2012; Felser *et al.*, 2012; Lim and Christianson, 2015; Boxell and Felser, 2017).

Table 4.11 Regions of interest in the volitional condition

		matrix predicate		pre-critical	critical	spillover	
subjunctive	Victoria	préfère	que	Claude	soit	responsable	du projet
indicative	Victoria	préfère	que	Claude	est	responsable	du projet

4.7 Results: Study 1a (Comprehension)

4.7.1 Data Preparation and Analyses

Following standard procedure in psycholinguistic research (Rayner and Pollatsek, 1989), fixations shorter than 80 ms were merged with the neighbouring fixation if the two fixations were within one character of each other. All other fixations shorter than 80 ms or exceeding 800 ms were excluded from the dataset (Rayner 1998; Rayner et al. 2004). Fixations occurring immediately before or after a blink were also removed from the analysis. Finally, we deleted fixations that were more than three standard deviations from the mean of each dependent

[19] In the experimental stimuli, the word following the critical region varied considerably in length. In order to minimise this variation and ensure a certain level of consistent across test items, we extended the spillover region to include up to three words. We recognise that the variation in the spillover region length is a confounding factor in this study and controlled for it post-hoc by including region length (in characters) as a control variable in the mixed-effects models.

variable. Comprehension question accuracy was high for both test groups (L1 $M = 0.91$, $SD = 0.06$; L2 $M = 0.89$, $SD = 0.06$), suggesting little to no impairment to offline comprehension.

To correct the skewed distribution and thus approximate normal distribution, fixation data were log-transformed. For continuous variables, mixed-effects linear regression models were conducted using the *lmer* function of the *lme4* package (Bates *et al.*, 2015) in the R environment (R Development Core Team, 2014). For binary variables[20] (regression probabilities), mixed-effects logistic regression models were computed using the *glmer* function of the aforementioned package.

Where possible, we fitted each model using the 'maximal' random effects structure (i.e., slopes for each fixed effect across item and subject) that converged (Barr *et al.*, 2013). In many cases, the maximal random effects structure had to be cut down due to non-convergence or a singular fit (i.e., perfect or near perfect correlations in the random structure). If a model did not converge or had a singular fit, parameters were systematically excluded from the random effects structure based on the level of variance. Parameters with the lowest variance were removed first until the model converged. Absolute t values (or z values in the case of binary variables) exceeding $\pm$ 1.96 were analysed as significant.

In light of our earlier discussion on the comparative fallacy in Section 4.6.1, L1 and L2 fixation data were analysed separately. For the L1 data, mixed-effect models were computed for each fixation measure to investigate whether there was a significant difference between mood (indicative and subjunctive), matrix semantic property (desire, directive and emotive-factive) and/or their interaction. Both mood and matrix semantic property were inputted as categorical fixed effects and sum contrasts were coded. For mood, we explored the difference between indicative and subjunctive sentences (indicative -1, subjunctive 1). For matrix semantic property, we examined the difference between desire and directive sentences (directive -1, desire 1, emotive-factive 0) and between emotive-factive and directive sentences (directive -1, desire 0,

[20] Research (e.g., Brysbaert, Drieghe and Vitu, 2005) has shown that word length is the strongest influence by far on skipping rates, meaning that even a difference of two letters can have a significant impact on skipping rate. The shorter length of the indicative forms when compared to the subjunctive forms are therefore likely to explain higher skipping rates with indicative than subjunctive sentences, regardless of grammaticality. It is therefore not possible to conduct skipping rate models without including worth length. For this reason, we only report the descriptive statistics for skipping probability.

emotive-factive 1).[21] Directive sentences were chosen as the baseline since these are the only verbs in English that allow the subjunctive mood.

To control for frequency and/or length confounds, we inputted at least one control variable as a continuous fixed effect. For global measures, we chose sentence length (centred) in characters and for local measures, region frequency or length (centred) in characters. We originally planned to include both region frequency and length for the critical region. However, this would have led to multicollinearity of variables between frequency and length. Consequently, we only selected region frequency in the critical region analyses and region length in the spillover region analyses. We also entered trial index as a continuous fixed factor to avoid any confounds relating to the position of each test item within the experiment. Although the addition of these parameters as control variables were important, we will not discuss them in any more depth. The mixed-effect models for the L2 data were almost identical to the L1 models. The only exception was the inclusion of proficiency (centred) as a continuous fixed effect and residence in a French-speaking country as a categorical fixed effect in the L2 models.

Finally, any significant fixed effects and/or interactions were interpreted based on estimated (least-squares) means, calculated using the emmeans package (Lenth, 2018), and predicted value plots. Least-square means refer to the group means that are adjusted in response to the means of other factors in the regression model. Predicted values concern the values that the regression model predicts for each case.

[21] Categorical variables with more than two factors require the use of dummy coding since we are unable to distinguish between factors with a single variable coded with zeros and ones (Field, Miles and Field, 2014).

Table 4.12 Raw means (standard deviations) for L1 global in comprehension

		Total Sentence Reading Time	Mean Fixation Count	Mean Fixation Duration
Desire	Indicative	3812.2	11.4	235.1
		(1602.5)	(5.3)	(34.9)
	Subjunctive	3350.6	9.9	230.6
		(1275.3)	(4.0)	(35.9)
Directive	Indicative	4412.1	13.1	240.9
		(1640.2)	(5.3)	(35.4)
	Subjunctive	4141.0	13.0	232.2
		(1682.0)	(5.7)	(31.8)
Emotive-Factive	Indicative	3984.7	11.8	235.3
		(1577.0)	(4.6)	(32.3)
	Subjunctive	3831.4	11.3	234.7
		(1615.0)	(4.7)	(37.3)

Table 4.13 Raw means (standard deviations) for the L2 global in comprehension

		Total Sentence Reading Time	Mean Fixation Count	Mean Fixation Duration
Desire	Indicative	5223.6	15.4	247.6
		(2235.1)	(6.7)	(35.3)
	Subjunctive	5280.9	15.4	251.8
		(2274.9)	(6.9)	(31.3)
Directive	Indicative	6181.3	18.1	253.4
		(2250.4)	(7.1)	(29.9)
	Subjunctive	6205.1	18.3	254
		(2268.6)	(7.1)	(32.2)
Emotive-Factive	Indicative	5294	15.5	250.7
		(2238.7)	(6.9)	(31.9)
	Subjunctive	5776.3	17	250.3
		(2372.7)	(7.2)	(31.6)

Table 4.14 Mixed-effects models for L1 global in comprehension

Measure	Fixed Effect	*b*	*SE*	*z*
Mean Fixation Count	(Intercept)	12.98	0.59	22.07
	Mood: Subjunctive vs. Indicative	-0.48	0.11	-4.37
	Semantic: Desire vs. Directive	-0.55	0.22	-2.45
	Semantic: Emotive-Factive vs. Directive	-0.15	0.22	-0.70
	Sentence Length (Centred)	1.92	0.16	11.92
	Trial Number	-0.03	0.01	-5.51
	Mood × Semantic (Desire vs. Directive)	-0.37	0.15	-2.42
	Mood × Semantic (Emotive-Factive vs. Directive)	0.10	0.15	0.68
Mean Fixation Duration	(Intercept)	5.46	0.02	309.38
	Mood: Subjunctive vs. Indicative	-0.01	0.00	-2.82
	Semantic: Desire vs. Directive	-0.01	0.01	-1.68
	Semantic: Emotive-Factive vs. Directive	0.00	0.01	0.08
	Sentence Length (Centred)	-0.01	0.00	-1.17
	Trial Number	0.00	0.00	-2.19
	Mood × Semantic (Desire vs. Directive)	0.00	0.00	0.05
	Mood × Semantic (Emotive-Factive vs. Directive)	0.01	0.00	1.58
Total Sentence Reading Time	(Intercept)	8.37	0.05	180.64
	Mood: Subjunctive vs. Indicative	-0.05	0.01	-6.33
	Semantic: Desire vs. Directive	-0.04	0.01	-2.95
	Semantic: Emotive-Factive vs. Directive	0.00	0.01	-0.24
	Sentence Length (Centred)	0.15	0.01	15.27
	Trial Number	0.00	0.00	-11.23
	Mood × Semantic (Desire vs. Directive)	-0.02	0.01	-1.61
	Mood × Semantic (Emotive-Factive vs. Directive)	0.02	0.01	1.73

Note: Shaded cells denote statistical significance ($p < 0.05$).

Table 4.15 Mixed-effects models for L2 global in comprehension

	Mean Fixation Count			Mean Fixation Duration			Total Sentence Reading Time		
	b	*SE*	*z*	*b*	*SE*	*z*	*b*	*SE*	*z*
(Intercept)	13.67	2.64	5.18	5.54	0.04	126.14	13.67	2.64	5.18
Mood: Subjunctive vs. Indicative	0.05	0.44	0.10	0.02	0.01	1.37	0.05	0.44	0.10
Semantic: Desire vs. Directive	-1.07	0.75	-1.41	-0.01	0.01	-0.65	-1.07	0.75	-1.41
Semantic: Emotive-Factive vs. Directive	-0.13	0.75	-0.18	0.00	0.01	0.14	-0.13	0.75	-0.18
Proficiency (Centred)	-2.78	2.78	-1.00	0.03	0.05	0.57	-2.78	2.78	-1.00
Residence (in French-Speaking Country)	3.92	2.94	1.33	-0.03	0.05	-0.68	3.92	2.94	1.33
Sentence Length (Centred)	1.91	0.33	5.85	0.00	0.00	-0.10	1.91	0.33	5.85
Trial Number	-0.03	0.01	-5.11	0.00	0.00	-0.89	-0.03	0.01	-5.11
Mood × Semantic (Desire vs. Directive)	-0.60	0.63	-0.95	0.00	0.01	-0.24	-0.60	0.63	-0.95
Mood × Semantic (Emotive-Factive vs. Directive)	1.20	0.63	1.93	0.03	0.01	2.35	1.20	0.63	1.93
Mood × Prof.	0.06	0.47	0.13	0.02	0.01	1.35	0.06	0.47	0.13
Semantic (Desire vs. Directive) × Prof.	-0.64	0.66	-0.97	0.00	0.01	-0.05	-0.64	0.66	-0.97
Semantic (Emotive-Factive vs. Directive) × Prof.	-0.20	0.66	-0.30	0.01	0.01	0.41	-0.20	0.66	-0.30
Mood × Residence	-0.03	0.50	-0.06	-0.02	0.01	-1.20	-0.03	0.50	-0.06
Semantic (Desire vs. Directive) × Residence	0.14	0.70	0.20	0.00	0.02	0.29	0.14	0.70	0.20
Semantic (Emotive-Factive vs. Directive) × Residence	-0.73	0.70	-1.05	0.00	0.02	-0.16	-0.73	0.70	-1.05
Proficiency × Residence	0.18	3.22	0.06	-0.07	0.05	-1.35	0.18	3.22	0.06
Mood × Semantic (Desire vs. Directive) × Prof.	-0.14	0.66	-0.22	0.00	0.01	-0.28	-0.14	0.66	-0.22
Mood × Semantic (Emotive-Factive vs. Directive) × Prof.	0.94	0.66	1.42	0.04	0.01	2.54	0.94	0.66	1.42
Mood × Semantic (Desire vs. Directive) × Residence	0.55	0.70	0.78	0.01	0.02	0.89	0.55	0.70	0.78
Mood × Semantic (Emotive-Factive vs. Directive) × Residence	-1.06	0.70	-1.51	-0.04	0.02	-2.55	-1.06	0.70	-1.51
Mood × Prof. × Residence	-0.53	0.54	-0.97	-0.02	0.01	-1.64	-0.53	0.54	-0.97
Semantic (Desire vs. Directive) × Prof. × Residence	0.33	0.76	0.43	0.00	0.02	0.13	0.33	0.76	0.43
Semantic (Emotive-Factive vs. Directive) × Prof. × Residence	-0.30	0.77	-0.39	0.00	0.02	-0.03	-0.30	0.77	-0.39
Mood × Semantic (Desire vs. Directive) × Prof. × Residence	0.09	0.77	0.12	0.00	0.02	-0.03	0.09	0.77	0.12
Mood × Semantic (Emotive-Factive vs. Directive) × Prof. × Residence	-1.17	0.77	-1.52	-0.04	0.02	-2.10	-1.17	0.77	-1.52

Note: Shaded cells denote statistical significance (p < 0.05

4.7.2 Global Measures

4.7.2.1 L1 Group

To examine the overall effect of mood and its interaction with matrix semantic property on reading behaviour, we calculated the following global measures: mean fixation duration, mean fixation count and total sentence reading time. For each measure, observed means are presented in Table 4.12 and model parameters in Table 4.14

All three global measures demonstrated a significant main effect of mood, with longer reading times (or in the case of mean fixation count, more fixations) for indicative than subjunctive sentences. There was a significant main effect of matrix semantic property (desire vs. directive) and a significant interaction between mood and matrix semantic property (desire vs. directive) for mean fixation count, suggesting that the size of the mood effect was stronger with desire than directive verbs.

4.7.2.2 L2 Group

To examine L2 learners' sensitivity to mood and the influence of matrix semantic property, proficiency and residence abroad, we calculated the following global measures: mean fixation duration, mean fixation count and total sentence reading time. For each measure, observed means are presented in Table 4.13 and model parameters in Table 4.15.

The descriptive statistics in Table 4.13 show that the raw mean fixation count, mean fixation duration and total sentence reading time were, on average, higher for subjunctive than indicative sentences. The only exception was mean fixation duration with emotive-factive verbs where it was marginally higher for indicative sentences.

Mixed-effects models, however, did not reveal a significant main effect of mood, or an interaction with any of the other predictor variables, for mean fixation count or total sentence reading time. The following significant interactions were observed, however, for mean fixation duration: Mood × Semantic (Emotive-Factive vs. Directive), Mood × Semantic (Emotive-Factive vs. Directive) × Proficiency and Mood × Semantic (Emotive-Factive vs. Directive) × Residence. In particular, mean fixation duration was longer for indicative than subjunctive sentences with directive verbs, whereas the reverse was true for emotive-factive verbs.

Proficiency and residence abroad, as individual fixed factors but not their interaction, further modulated this effect. Figure 4.3 shows that although at lower proficiency levels, mean fixation duration was almost identical for indicative and subjunctive complements to directive predicates, the difference between the two complement types increased as a function of proficiency, such that mean fixation duration became noticeably longer, but not at a statistically significant level, for indicative than subjunctive sentences with directive verbs. However, mean fixation duration was, for the most part but again not at a statistically significant level, longer for subjunctive than indicative complements to desire[22] and emotive-factive verbs.

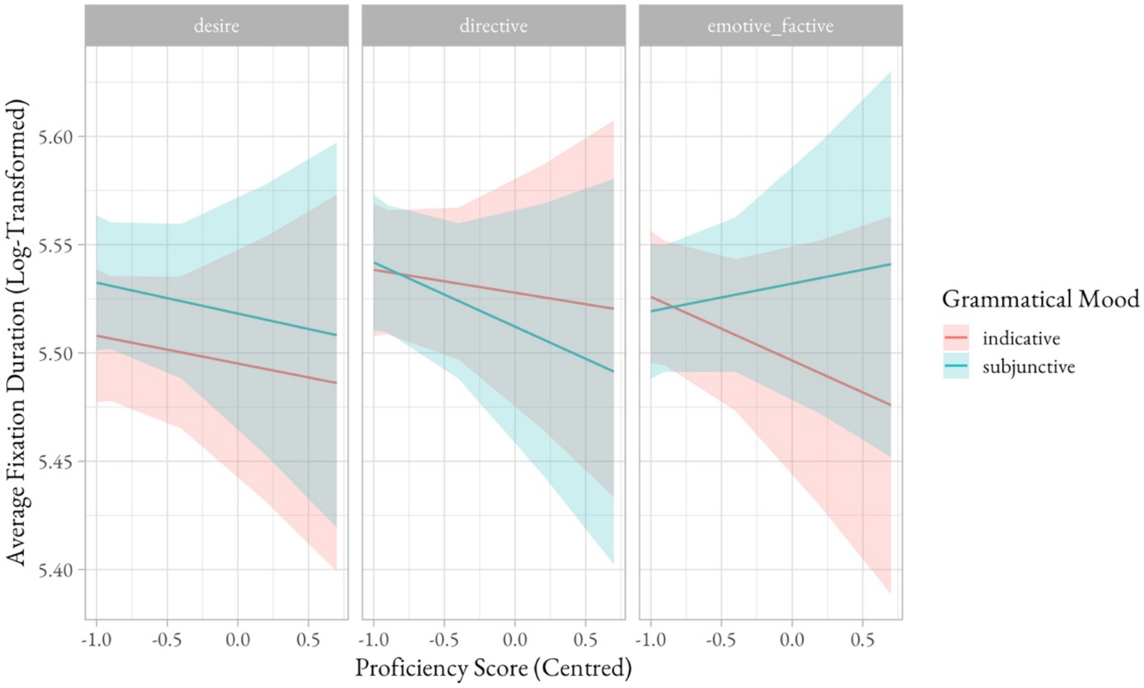


Figure 4.3 Predicted mean fixation duration (comprehension)

Finally, mean fixation duration was longer for indicative than subjunctive sentences with directive verbs among both the residence and non-residence abroad group, but the effect was stronger among the non-residence abroad group. Mean fixation duration among the residence abroad group was longer for indicative than subjunctive sentences with emotive-factive verbs, but longer for subjunctive than indicative sentences among the non-residence abroad group.

[22] Note, however, that the difference between desire and directive verbs was not significant.

Table 4.16 Raw means (standard deviations) for L1 local in comprehension

		First Fixation Duration	Gaze Duration	Go-Past Time	Total Reading Time	Skipping Probability	Regression-In Probability	Regression-Out Probability
Critical Region								
Desire	Indicative	256.25 (116.47)	272.31 (143.99)	274.14 (152.6)	372.97 (255.24)	0.73 (0.45)	0.62 (0.49)	0.03 (0.17)
	Subjunctive	226.05 (59.21)	230.03 (63.96)	233.1 (65.73)	279.82 (150.79)	0.62 (0.49)	0.17 (0.37)	0.08 (0.27)
Directive	Indicative	250.81 (97.58)	272.96 (133.75)	278.39 (137.86)	360.99 (206.52)	0.67 (0.47)	0.44 (0.5)	0.06 (0.24)
	Subjunctive	242.43 (81.4)	257.63 (101.68)	259.17 (102.12)	320.7 (157.79)	0.47 (0.5)	0.16 (0.37)	0.04 (0.18)
Emotive-Factive	Indicative	257.55 (102.8)	269.06 (115.15)	275.14 (129.3)	330.09 (184.75)	0.65 (0.48)	0.47 (0.5)	0.05 (0.22)
	Subjunctive	243.71 (91.26)	255.05 (105.01)	266.03 (126.49)	319.11 (182.9)	0.51 (0.5)	0.17 (0.38)	0.09 (0.29)
Spillover Region								
Desire	Indicative	242.42 (86.38)	406.18 (232.56)	511.31 (309.96)	662.7 (407.4)	0.11 (0.32)	0.23 (0.42)	0.41 (0.49)
	Subjunctive	243.3 (73.31)	398.73 (230.3)	426.96 (241.36)	513.76 (292.77)	0.08 (0.27)	0.15 (0.36)	0.21 (0.41)
Directive	Indicative	249.58 (97)	401.14 (231.71)	498.23 (300.01)	669.49 (414.32)	0.11 (0.32)	0.24 (0.43)	0.29 (0.45)
	Subjunctive	230.35 (77.12)	335.4 (177.17)	368.2 (210.71)	501.95 (312.03)	0.13 (0.34)	0.24 (0.43)	0.17 (0.37)
Emotive-Factive	Indicative	240.9 (75.43)	360.22 (219.17)	420.12 (242.24)	577.22 (346.29)	0.16 (0.37)	0.22 (0.42)	0.32 (0.47)
	Subjunctive	246.49 (86.6)	356.64 (219.63)	383.97 (244.16)	481.64 (302.23)	0.18 (0.38)	0.15 (0.36)	0.17 (0.38)

Table 4.17 Raw means (standard deviations) for L2 local in comprehension

		First Fixation Duration	Gaze Duration	Go-Past Time	Total Reading Time	Skipping Probability	Regression-In Probability	Regression-Out Probability
				Critical Region				
Desire	Indicative	257.34 (97.73)	264.63 (108.85)	274.61 (120.91)	378.85 (250.64)	0.62 (0.49)	0.44 (0.5)	0.06 (0.24)
	Subjunctive	273.65 (107.56)	286.41 (131.96)	294.08 (143.37)	429.48 (259.03)	0.48 (0.5)	0.3 (0.46)	0.08 (0.28)
Directive	Indicative	255.49 (96.18)	260.75 (110.84)	274.96 (131.28)	401.13 (275.51)	0.62 (0.49)	0.46 (0.5)	0.09 (0.28)
	Subjunctive	274.61 (94.93)	296.33 (116.8)	306.21 (121.77)	451.63 (241.27)	0.39 (0.49)	0.26 (0.44)	0.06 (0.23)
Emotive-Factive	Indicative	245.53 (82.52)	259.16 (106.66)	274.13 (127.14)	377.85 (227.98)	0.58 (0.49)	0.35 (0.48)	0.13 (0.34)
	Subjunctive	261.92 (95.08)	278.03 (127.92)	294.92 (155.81)	444.68 (305.1)	0.4 (0.49)	0.33 (0.47)	0.12 (0.33)
				Spillover Region				
Desire	Indicative	268.20 (99.71)	476.27 (312.69)	595.51 (357.39)	907.55 (580.94)	0.07 (0.26)	0.38 (0.49)	0.3 (0.46)
	Subjunctive	268.77 (97.57)	501.93 (257.48)	588.08 (293.33)	857.81 (506.79)	0.06 (0.24)	0.34 (0.48)	0.21 (0.41)
Directive	Indicative	268.09 (100.11)	460.47 (280)	575.13 (401.91)	964.59 (649.27)	0.08 (0.28)	0.41 (0.49)	0.23 (0.42)
	Subjunctive	276.01 (100.99)	456.08 (241.64)	525.6 (330.33)	848.68 (546.64)	0.06 (0.24)	0.36 (0.48)	0.17 (0.38)
Emotive-Factive	Indicative	266.02 (102.64)	413.42 (269.94)	499.09 (316.46)	764.65 (544.02)	0.13 (0.33)	0.31 (0.46)	0.24 (0.43)
	Subjunctive	265.22 (96.42)	464.18 (330.14)	576.7 (396.22)	848.34 (566.98)	0.07 (0.25)	0.25 (0.43)	0.29 (0.45)

Table 4.18 Mixed-effects models for L1 local in comprehension

		Critical Region			Spillover Region		
		b	*SE*	*t/z*	*b*	*SE*	*t/z*
First Fixation Duration	(Intercept)	5.41	0.03	155.26	5.44	0.03	206.37
	Mood: Subjunctive vs. Indicative	-0.08	0.04	-2.03	-0.01	0.01	-0.79
	Semantic: Desire vs. Directive	-0.02	0.02	-0.85	0.01	0.02	0.36
	Semantic: Emotive-Factive vs. Directive	0.02	0.02	1.15	0.01	0.02	0.30
	Region Frequency (Centred)	-0.06	0.04	-1.42	0.01	0.01	0.76
	Trial Number	0.00	0.00	0.92	0.00	0.00	0.08
	Mood × Semantic (Desire vs. Directive)	-0.03	0.02	-1.47	0.01	0.01	1.13
	Mood × Semantic (Emotive-Factive vs. Directive)	0.01	0.02	0.31	0.01	0.01	1.03
Gaze Duration	(Intercept)	5.44	0.04	141.24	5.80	0.04	137.06
	Mood: Subjunctive vs. Indicative	-0.09	0.05	-1.87	-0.02	0.02	-1.32
	Semantic: Desire vs. Directive	-0.03	0.02	-1.25	0.07	0.03	2.49
	Semantic: Emotive-Factive vs. Directive	0.02	0.02	0.81	-0.07	0.03	-2.31
	Region Frequency (Centred)	-0.06	0.04	-1.47	0.17	0.02	8.74
	Trial Number	0.00	0.00	1.20	0.00	0.00	-0.97
	Mood × Semantic (Desire vs. Directive)	-0.03	0.02	-1.40	0.02	0.03	0.70
	Mood × Semantic (Emotive-Factive vs. Directive)	0.01	0.02	0.51	0.02	0.03	0.90
Go-Past Time	(Intercept)	5.47	0.04	136.80	5.98	0.05	125.37
	Mood: Subjunctive vs. Indicative	-0.08	0.05	-1.61	-0.08	0.01	-5.70
	Semantic: Desire vs. Directive	-0.03	0.02	-1.32	0.09	0.04	2.13
	Semantic: Emotive-Factive vs. Directive	0.02	0.02	1.04	-0.09	0.04	-2.21
	Region Frequency (Centred)	-0.06	0.05	-1.29	0.21	0.03	7.17
	Trial Index	0.00	0.00	0.56	0.00	0.00	-2.96
	Mood × Semantic (Desire vs. Directive)	-0.03	0.02	-1.36	0.01	0.02	0.49
	Mood × Semantic (Emotive-Factive vs. Directive)	0.02	0.02	0.89	0.04	0.02	1.90
Total Reading Time	(Intercept)	5.69	0.05	113.08	6.26	0.06	110.33
	Mood: Subjunctive vs. Indicative	-0.10	0.05	-1.80	-0.11	0.02	-6.99
	Semantic: Desire vs. Directive	-0.03	0.03	-1.05	0.06	0.04	1.69
	Semantic: Emotive-Factive vs. Directive	-0.01	0.03	-0.29	-0.09	0.04	-2.58
	Region Frequency (Centred)	-0.05	0.05	-0.91	0.22	0.03	8.38
	Trial Index	0.00	0.00	-2.01	0.00	0.00	-4.40
	Mood × Semantic (Desire vs. Directive)	-0.06	0.03	-2.18	0.00	0.02	-0.08
	Mood × Semantic (Emotive-Factive vs. Directive)	0.04	0.03	1.53	0.03	0.02	1.18
Regression-In Probability	(Intercept)	-0.48	0.22	-2.20	-1.14	0.19	-5.94
	Mood: Subjunctive vs. Indicative	-0.94	0.27	-3.46	-0.16	0.08	-1.97
	Semantic: Desire vs. Directive	0.23	0.15	1.54	-0.08	0.12	-0.71
	Semantic: Emotive-Factive vs. Directive	-0.06	0.14	-0.40	-0.15	0.12	-1.26
	Region Frequency (Centred)	-0.06	0.26	-0.24	-0.02	0.08	-0.25
	Trial Index	-0.01	0.00	-2.07	-0.01	0.00	-2.56
	Mood × Semantic (Desire vs. Directive)	-0.20	0.15	-1.37	-0.09	0.12	-0.80
	Mood × Semantic (Emotive-Factive vs. Directive)	0.10	0.14	0.72	-0.06	0.12	-0.52
Regression-Out Probability	(Intercept)	-2.84	0.36	-7.97	-0.97	0.15	-6.58
	Mood: Subjunctive vs. Indicative	0.04	0.51	0.07	-0.41	0.08	-5.37
	Semantic: Desire vs. Directive	-0.11	0.28	-0.39	0.24	0.10	2.34
	Semantic: Emotive-Factive vs. Directive	0.27	0.26	1.07	-0.08	0.11	-0.74
	Region Frequency (Centred)	-0.16	0.50	-0.32	0.17	0.07	2.29
	Trial Index	0.00	0.01	-0.16	0.00	0.00	-1.10
	Mood × Semantic (Desire vs. Directive)	0.31	0.28	1.12	-0.08	0.10	-0.74
	Mood × Semantic (Emotive-Factive vs. Directive)	0.18	0.25	0.71	0.02	0.11	0.14

Note: Shaded cells denote statistical significance ($p < 0.05$).

Table 4.19 Mixed-effects models for L2 local in comprehension

		Critical Region			Spillover Region		
		b	*SE*	*t/z*	*b*	*SE*	*t/z*
First Fixation Duration	(Intercept)	5.49	0.07	75.35	5.58	0.06	90.85
	Mood: Subjunctive vs. Indicative	0.00	0.05	-0.01	0.03	0.03	0.94
	Semantic: Desire vs. Directive	0.01	0.06	0.12	-0.07	0.05	-1.58
	Semantic: Emotive-Factive vs. Directive	0.07	0.06	1.21	-0.01	0.05	-0.25
	Proficiency (Centred)	0.02	0.07	0.22	0.03	0.06	0.46
	Residence (in French-Speaking Country)	-0.01	0.08	-0.08	-0.02	0.07	-0.37
	Region Frequency / Length (Centred)	-0.08	0.04	-2.27	0.01	0.01	0.50
	Trial Number	0.00	0.00	0.89	0.00	0.00	-1.60
	Mood × Semantic (Desire vs. Directive)	-0.01	0.06	-0.16	0.02	0.05	0.34
	Mood × Semantic (Emotive-Factive vs. Directive)	-0.04	0.06	-0.66	0.02	0.05	0.46
	Mood × Prof.	0.04	0.04	0.94	0.03	0.03	1.03
	Semantic (Desire vs. Directive) × Prof.	-0.02	0.06	-0.34	-0.08	0.05	-1.60
	Semantic (Emotive-Factive vs. Directive) × Prof.	0.10	0.06	1.63	0.01	0.05	0.11
	Mood × Residence	-0.06	0.05	-1.37	-0.04	0.04	-1.01
	Semantic (Desire vs. Directive) × Residence	0.02	0.06	0.31	0.09	0.05	1.83
	Semantic (Emotive-Factive vs. Directive) × Residence	-0.10	0.06	-1.63	0.01	0.05	0.29
	Proficiency × Residence	-0.05	0.08	-0.58	-0.04	0.07	-0.59
	Mood × Semantic (Desire vs. Directive) × Prof.	0.05	0.06	0.87	0.04	0.05	0.87
	Mood × Semantic (Emotive-Factive vs. Directive) × Prof.	-0.03	0.06	-0.48	0.00	0.05	-0.03
	Mood × Semantic (Desire vs. Directive) × Residence	0.04	0.06	0.62	0.00	0.05	-0.06
	Mood × Semantic (Emotive-Factive vs. Directive) × Residence	0.03	0.06	0.44	-0.04	0.05	-0.79
	Mood × Prof. × Residence	-0.06	0.05	-1.30	-0.06	0.04	-1.58
	Semantic (Desire vs. Directive) × Prof. × Residence	0.05	0.07	0.69	0.09	0.05	1.62
	Semantic (Emotive-Factive vs. Directive) × Prof. × Residence	-0.12	0.07	-1.81	0.02	0.06	0.31
	Mood × Semantic (Desire vs. Directive) × Prof. × Residence	-0.05	0.07	-0.70	-0.03	0.06	-0.63
	Mood × Semantic (Emotive-Factive.) × Prof. × Residence	0.01	0.07	0.15	0.02	0.06	0.39
Gaze Duration	(Intercept)	5.50	0.08	66.36	6.12	0.09	66.66
	Mood: Subjunctive vs. Indicative	0.00	0.06	0.00	0.05	0.05	0.90
	Semantic: Desire vs. Directive	0.01	0.07	0.20	0.08	0.08	1.01
	Semantic: Emotive-Factive vs. Directive	0.04	0.06	0.69	-0.15	0.08	-1.86
	Proficiency (Centred)	0.00	0.08	-0.04	0.09	0.09	1.03
	Residence (in French-Speaking Country)	0.01	0.09	0.16	-0.15	0.10	-1.60
	Region Frequency / Length (Centred)	-0.07	0.04	-1.64	0.17	0.03	6.10
	Trial Number	0.00	0.00	0.43	0.00	0.00	-2.79
	Mood × Semantic (Desire vs. Directive)	0.01	0.06	0.23	-0.06	0.07	-0.77
	Mood × Semantic (Emotive-Factive vs. Directive)	-0.06	0.06	-1.01	0.14	0.07	1.85
	Mood × Prof.	0.00	0.05	0.07	0.03	0.05	0.50
	Semantic (Desire vs. Directive) × Prof.	0.00	0.07	-0.01	0.00	0.08	-0.05
	Semantic (Emotive-Factive vs. Directive) × Prof.	0.06	0.06	0.95	-0.07	0.08	-0.86
	Mood × Residence	-0.03	0.05	-0.67	0.00	0.06	-0.01
	Semantic (Desire vs. Directive) × Residence	0.00	0.07	0.07	0.00	0.08	-0.06
	Semantic (Emotive-Factive vs. Directive) × Residence	-0.08	0.07	-1.11	0.05	0.08	0.65
	Proficiency × Residence	-0.05	0.10	-0.55	-0.19	0.10	-1.83
	Mood × Semantic (Desire vs. Directive) × Prof.	0.08	0.07	1.24	-0.06	0.08	-0.81
	Mood × Semantic (Emotive-Factive vs. Directive) × Prof.	-0.06	0.06	-1.01	0.13	0.08	1.67
	Mood × Semantic (Desire vs. Directive) × Residence	0.01	0.07	0.13	0.09	0.08	1.08
	Mood × Semantic (Emotive-Factive vs. Directive) × Residence	0.04	0.07	0.52	-0.14	0.08	-1.73
	Mood × Prof. × Residence	-0.01	0.05	-0.20	0.00	0.06	-0.05
	Semantic (Desire vs. Directive) × Prof. × Residence	0.03	0.07	0.35	0.02	0.09	0.24
	Semantic (Emotive-Factive vs. Directive) × Prof. × Residence	-0.09	0.07	-1.22	0.04	0.09	0.40
	Mood × Semantic (Desire vs. Directive) × Prof. × Residence	-0.08	0.08	-1.01	0.08	0.09	0.90
	Mood × Semantic (Emotive-Factive vs. Directive) × Prof. × Residence	0.05	0.07	0.69	-0.11	0.09	-1.23
Go-Past Time	(Intercept)	5.52	0.09	61.10	6.21	0.10	60.59
	Mood: Subjunctive vs. Indicative	-0.02	0.06	-0.31	0.04	0.05	0.85
	Semantic: Desire vs. Directive	0.01	0.07	0.16	0.05	0.08	0.63
	Semantic: Emotive-Factive vs. Directive	0.06	0.07	0.94	-0.16	0.08	-1.97
	Proficiency (Centred)	0.00	0.09	-0.04	0.06	0.10	0.63
	Residence (in French-Speaking Country)	0.04	0.10	0.42	-0.11	0.11	-0.98

		Critical Region			Spillover Region		
		b	*SE*	*t/z*	*b*	*SE*	*t/z*
	Region Frequency / Length (Centred)	-0.08	0.04	-2.04	0.23	0.03	6.76
	Trial Number	0.00	0.00	0.19	0.00	0.00	-1.87
	Mood × Semantic (Desire vs. Directive)	0.01	0.07	0.15	-0.12	0.07	-1.81
	Mood × Semantic (Emotive-Factive vs. Directive)	-0.07	0.07	-1.08	0.26	0.07	3.83
	Mood × Prof.	0.00	0.05	-0.05	0.02	0.05	0.45
	Semantic (Desire vs. Directive) × Prof.	0.01	0.07	0.14	-0.04	0.07	-0.53
	Semantic (Emotive-Factive vs. Directive) × Prof.	0.06	0.07	0.90	-0.08	0.07	-1.11
	Mood × Residence	-0.06	0.05	-1.08	-0.01	0.05	-0.22
	Semantic (Desire vs. Directive) × Residence	0.00	0.07	-0.01	0.05	0.08	0.71
	Semantic (Emotive-Factive vs. Directive) × Residence	-0.10	0.07	-1.35	0.07	0.08	0.92
	Proficiency × Residence	-0.05	0.11	-0.43	-0.25	0.12	-2.11
	Mood × Semantic (Desire vs. Directive) × Prof.	0.09	0.07	1.22	-0.12	0.07	-1.72
	Mood × Semantic (Emotive-Factive vs. Directive) × Prof.	-0.08	0.07	-1.15	0.23	0.07	3.17
	Mood × Semantic (Desire vs. Directive) × Residence	0.02	0.08	0.21	0.15	0.08	1.94
	Mood × Semantic (Emotive-Factive vs. Directive) × Residence	0.05	0.07	0.68	-0.30	0.08	-3.86
	Mood × Prof. × Residence	-0.05	0.06	-0.83	0.00	0.06	-0.05
	Semantic (Desire vs. Directive) × Prof. × Residence	0.01	0.08	0.15	0.05	0.08	0.55
	Semantic (Emotive-Factive vs. Directive) × Prof. × Residence	-0.10	0.08	-1.26	0.04	0.08	0.50
	Mood × Semantic (Desire vs. Directive) × Prof. × Residence	-0.08	0.08	-0.95	0.16	0.08	1.92
	Mood × Semantic (Emotive-Factive vs. Directive) × Prof. × Residence	0.07	0.08	0.84	-0.29	0.08	-3.42
Total Reading Time	(Intercept)	5.54	0.14	40.85	6.50	0.16	40.73
	Mood: Subjunctive vs. Indicative	-0.08	0.10	-0.86	0.03	0.05	0.57
	Semantic: Desire vs. Directive	-0.03	0.10	-0.30	-0.04	0.09	-0.39
	Semantic: Emotive-Factive vs. Directive	0.08	0.10	0.79	-0.16	0.09	-1.71
	Proficiency (Centred)	-0.15	0.14	-1.11	-0.04	0.16	-0.22
	Residence (in French-Speaking Country)	0.27	0.15	1.85	0.06	0.17	0.38
	Region Frequency / Length (Centred)	-0.12	0.08	-1.59	0.25	0.04	5.54
	Trial Number	0.00	0.00	1.67	0.00	0.00	-3.67
	Mood × Semantic (Desire vs. Directive)	-0.20	0.10	-2.07	-0.06	0.07	-0.79
	Mood × Semantic (Emotive-Factive vs. Directive)	0.06	0.09	0.68	0.22	0.07	2.98
	Mood × Prof.	-0.07	0.07	-1.00	0.02	0.05	0.44
	Semantic (Desire vs. Directive) × Prof.	0.02	0.10	0.18	-0.08	0.08	-1.07
	Semantic (Emotive-Factive vs. Directive) × Prof.	0.05	0.10	0.51	-0.09	0.08	-1.18
	Mood × Residence	0.01	0.07	0.13	-0.04	0.06	-0.70
	Semantic (Desire vs. Directive) × Residence	0.02	0.10	0.15	0.08	0.08	1.03
	Semantic (Emotive-Factive vs. Directive) × Residence	-0.11	0.10	-1.06	0.01	0.08	0.12
	Proficiency × Residence	0.06	0.16	0.41	-0.17	0.19	-0.92
	Mood × Semantic (Desire vs. Directive) × Prof.	-0.15	0.10	-1.49	-0.04	0.08	-0.57
	Mood × Semantic (Emotive-Factive vs. Directive) × Prof.	0.09	0.10	0.94	0.18	0.08	2.39
	Mood × Semantic (Desire vs. Directive) × Residence	0.19	0.11	1.80	0.08	0.08	0.94
	Mood × Semantic (Emotive-Factive vs. Directive) × Residence	-0.08	0.10	-0.82	-0.21	0.08	-2.57
	Mood × Prof. × Residence	0.02	0.08	0.24	-0.06	0.06	-0.93
	Semantic (Desire vs. Directive) × Prof. × Residence	-0.04	0.11	-0.35	0.04	0.09	0.49
	Semantic (Emotive-Factive vs. Directive) × Prof. × Residence	-0.07	0.11	-0.65	0.05	0.09	0.60
	Mood × Semantic (Desire vs. Directive) × Prof. × Residence	0.09	0.11	0.79	0.09	0.09	1.06
	Mood × Semantic (Emotive-Factive vs. Directive) × Prof. × Residence	-0.08	0.11	-0.76	-0.21	0.09	-2.36
Regression-In Probability	(Intercept)	-1.43	0.34	-4.23	-0.97	0.25	-3.90
	Mood: Subjunctive vs. Indicative	-0.42	0.36	-1.18	0.00	0.23	0.02
	Semantic: Desire vs. Directive	-0.66	0.47	-1.39	-0.11	0.33	-0.34
	Semantic: Emotive-Factive vs. Directive	0.15	0.43	0.36	-0.29	0.34	-0.84
	Proficiency (Centred)	-0.37	0.32	-1.17	-0.38	0.24	-1.59
	Residence (in French-Speaking Country)	0.80	0.34	2.37	0.43	0.26	1.67
	Region Frequency / Length (Centred)	0.15	0.19	0.77	0.10	0.05	1.84
	Trial Number	0.00	0.00	1.22	-0.01	0.00	-2.32
	Mood × Semantic (Desire vs. Directive)	-0.69	0.47	-1.46	0.62	0.33	1.89
	Mood × Semantic (Emotive-Factive vs. Directive)	0.73	0.43	1.72	-0.75	0.34	-2.21
	Mood × Prof.	-0.32	0.32	-1.01	0.12	0.24	0.49
	Semantic (Desire vs. Directive) × Prof.	-0.62	0.47	-1.31	-0.17	0.34	-0.52
	Semantic (Emotive-Factive vs. Directive) × Prof.	0.24	0.44	0.56	-0.13	0.35	-0.37
	Mood × Residence	0.30	0.34	0.88	-0.20	0.26	-0.77

		Critical Region			Spillover Region		
		b	*SE*	*t/z*	*b*	*SE*	*t/z*
	Semantic (Desire vs. Directive) × Residence	0.63	0.50	1.24	0.21	0.36	0.58
	Semantic (Emotive-Factive vs. Directive) × Residence	-0.14	0.46	-0.30	-0.01	0.38	-0.03
	Proficiency × Residence	0.22	0.35	0.62	0.05	0.28	0.17
	Mood × Semantic (Desire vs. Directive) × Prof.	-0.84	0.47	-1.80	0.65	0.34	1.91
	Mood × Semantic (Emotive-Factive vs. Directive) × Prof.	0.76	0.44	1.73	-0.81	0.35	-2.31
	Mood × Semantic (Desire vs. Directive) × Residence	0.56	0.51	1.12	-0.63	0.36	-1.73
	Mood × Semantic (Emotive-Factive vs. Directive) × Residence	-0.45	0.46	-0.98	0.81	0.38	2.13
	Mood × Prof. × Residence	0.31	0.35	0.88	-0.21	0.28	-0.77
	Semantic (Desire vs. Directive) * Prof. × Residence	0.37	0.52	0.71	0.10	0.39	0.25
	Semantic (Emotive-Factive vs. Directive) × Prof. × Residence	-0.08	0.49	-0.16	0.28	0.41	0.69
	Mood × Semantic (Desire vs. Directive) × Prof. × Residence	0.87	0.52	1.68	-0.69	0.39	-1.79
	Mood × Semantic (Emotive-Factive vs. Directive) × Prof. × Residence	-0.84	0.49	-1.74	0.96	0.40	2.36
Regression-Out Probability	(Intercept)	-2.67	0.61	-4.37	-2.09	0.30	-6.94
	Mood: Subjunctive vs. Indicative	-0.13	0.64	-0.21	-0.03	0.28	-0.11
	Semantic: Desire vs. Directive	0.24	0.83	0.29	-0.67	0.41	-1.64
	Semantic: Emotive-Factive vs. Directive	0.49	0.70	0.69	-0.09	0.40	-0.21
	Proficiency (Centred)	-0.13	0.60	-0.22	-0.40	0.28	-1.40
	Residence (in French-Speaking Country)	0.26	0.62	0.42	0.66	0.30	2.16
	Region Frequency / Length (Centred)	-0.05	0.31	-0.15	0.26	0.06	4.30
	Trial Number	0.00	0.01	-0.56	0.01	0.00	2.37
	Mood × Semantic (Desire vs. Directive)	-0.03	0.83	-0.04	-0.80	0.41	-1.96
	Mood × Semantic (Emotive-Factive vs. Directive)	-0.59	0.70	-0.84	0.76	0.40	1.89
	Mood × Prof.	0.12	0.60	0.19	0.00	0.28	0.01
	Semantic (Desire vs. Directive) × Prof.	0.59	0.88	0.67	-0.71	0.41	-1.74
	Semantic (Emotive-Factive vs. Directive) × Prof.	-0.40	0.71	-0.56	-0.10	0.40	-0.24
	Mood × Residence	0.06	0.62	0.09	-0.13	0.31	-0.43
	Semantic (Desire vs. Directive) × Residence	-0.12	0.89	-0.14	1.00	0.44	2.25
	Semantic (Emotive-Factive vs. Directive) × Residence	-0.64	0.78	-0.82	0.24	0.44	0.56
	Proficiency × Residence	-0.19	0.66	-0.29	0.20	0.32	0.61
	Mood × Semantic (Desire vs. Directive) × Prof.	0.01	0.88	0.02	-0.79	0.41	-1.95
	Mood × Semantic (Emotive-Factive vs. Directive) × Prof.	-0.65	0.71	-0.92	0.52	0.40	1.28
	Mood × Semantic (Desire vs. Directive) × Residence	-0.03	0.89	-0.04	0.57	0.44	1.28
	Mood × Semantic (Emotive-Factive vs. Directive) × Residence	0.72	0.78	0.92	-0.78	0.44	-1.78
	Mood × Prof. × Residence	-0.25	0.66	-0.39	0.00	0.32	-0.01
	Semantic (Desire vs. Directive) × Prof. × Residence	-0.18	0.96	-0.19	0.98	0.46	2.13
	Semantic (Emotive-Factive vs. Directive) × Prof. × Residence	-0.14	0.81	-0.18	0.01	0.46	0.03
	Mood × Semantic (Desire vs. Directive) × Prof. × Residence	-0.57	0.96	-0.60	0.73	0.46	1.61
	Mood × Semantic (Emotive-Factive vs. Directive) × Prof. × Residence	0.87	0.81	1.07	-0.88	0.46	-1.92

Note: Shaded cells denote statistical significance ($p < 0.05$)

4.7.3 Local Measures

4.7.3.1 L1 Group

To examine the overall effect of mood and its interaction with matrix semantic property on reading behaviour, we calculated the following local measures at the critical and spillover regions: first fixation duration, gaze duration, go-past time, total reading time, regression-in probability and regression-out probability. For a more in-depth description of these measures, see Section 2.2.4.1. The means (and standard deviations) and the fixed effect estimates for the local measures in the critical and spillover region are shown in Table 4.16 and Table 4.18, respectively. For sensitivity to be attested, we expected reading measures to be higher for indicative than subjunctive sentences.

In the critical region, we only observed a significant main effect of mood for first fixation duration, such that reading times were longer for indicative (than subjunctive) complements, regardless of matrix semantic property. In the spillover region, there was a significant main effect of mood for go-past time, total reading time, regression-in probability and regression-out probability. In particular, these measures were higher for indicative (than subjunctive) complements, regardless of matrix semantic property.

4.7.3.2 L2 Group

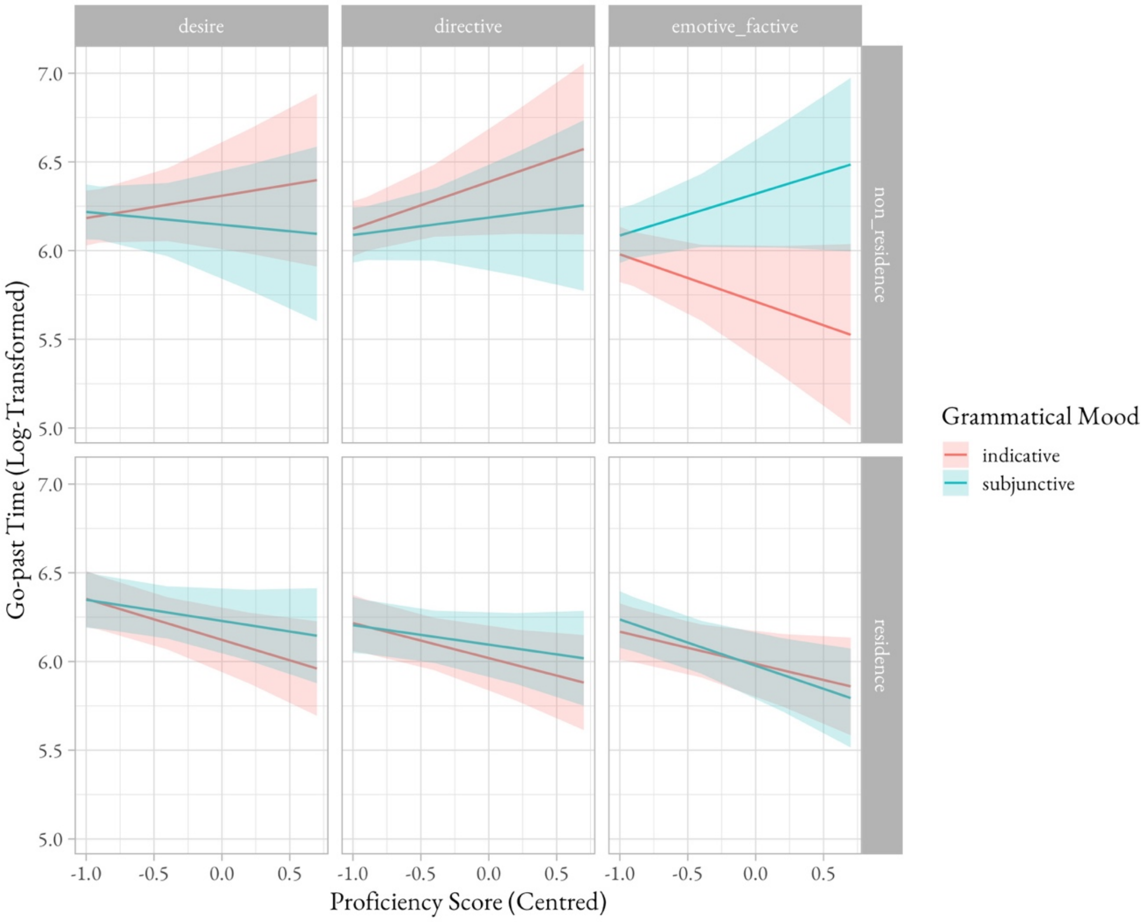
To examine L2 learners' sensitivity to mood and the influence of matrix semantic property, proficiency and residence abroad, we calculated the following local measures: first fixation duration, gaze duration, go-past time, total reading time, regression-in probability and regression-out probability. For a more in-depth description of these measures, see Section 2.2.4.1. The means (and standard deviations) and the fixed effect estimates for the local measures in the critical and spillover region are shown in Table 4.17 and Table 4.19, respectively.

In the critical region, we did not detect a significant main effect of mood for any of the reading time measures. A visual inspection of the raw data, however, revealed that reading times were numerically higher for subjunctive (than indicative) complements across each matrix semantic property.

There was, however, a significant two-way interaction between condition and matrix semantic property (desire vs. directive) for total reading time. In particular, total reading time was longer for subjunctive (than indicative) sentences with directive verbs, but the reverse was true for emotive-factive and desire verbs.

In the spillover region, mood alone did not predict any of the reading time measures. There were, however, several significant interactions. For each measure where a significant two-way, three-way or four-way interaction was observed, we only report the highest order interaction since this supersedes the significance of any previously reported lower order interactions.

Figure 4.4 Predicted go-past time in the spillover region (comprehension)



For go-past time, there was a reliable four-way interaction between condition, verb type (emotive-factive vs. directive), proficiency and residence abroad. Figure 4.4 visualises this interaction. Among the non-residence abroad group, go-past times were longer for indicative than subjunctive complements to directive (and desire) verbs especially as proficiency increased, but not at a statistically significant level. However, subjunctive complements to emotive-factive verbs attracted statistically significant longer go-past times than indicative complements, especially as proficiency increased. Among the residence abroad group, the difference between indicative and subjunctive complements was minimal, particularly with emotive-factive predicates. At lower proficiency levels, the difference between indicative and subjunctive complements to desire and directive verbs was negligible, but as proficiency increased, subjunctive complements attracted marginally longer go-past times.

For total reading time, there was a significant four-way interaction between condition, matrix semantic property (emotive-factive vs. directive), proficiency and residence abroad.

Figure 4.5 visualises this interaction. Total reading time was longer (but not at a statistically significant level) for indicative than subjunctive complements to directive verbs among both the non-residence abroad and residence abroad groups, but this difference was stronger among the non-residence abroad, especially as proficiency increased. Subjunctive complements to emotive-factive verbs attracted statistically significant longer reading times than indicative complements among the non-residence abroad group, particularly at higher proficiency levels. In contrast, among the residence abroad, total reading time was marginally higher for subjunctive complements than indicative complements to emotive-factive verbs, but this trend reversed as a function of proficiency. The difference in total reading time between indicative and subjunctive complements to desire verbs was minimal, regardless of residence abroad experience.

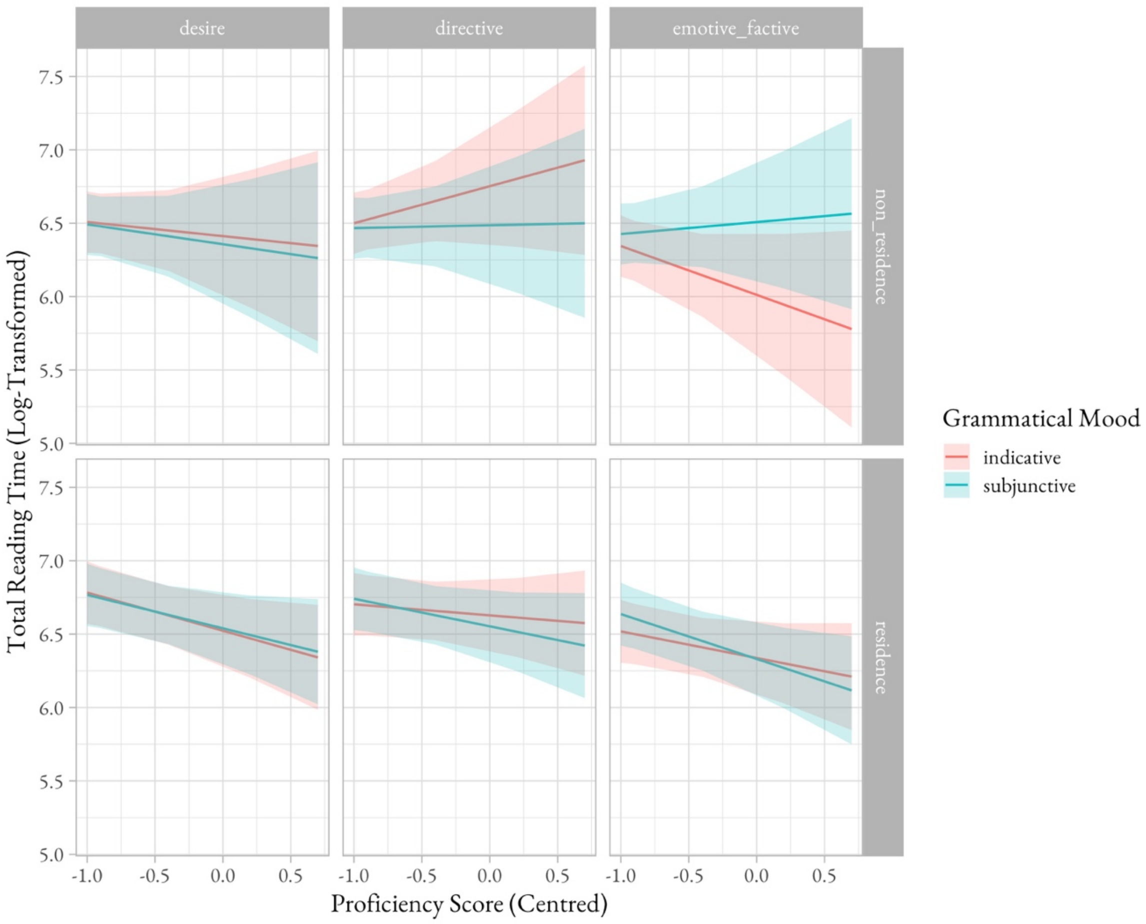


Figure 4.5 Predicted total reading time in the spillover region (comprehension)

For regression-in probability, there was a significant four-way interaction between condition, matrix semantic property (emotive-factive vs. directive), proficiency and residence abroad. Figure 4.6 shows that regression-in probability was numerically but not significantly higher for indicative than subjunctive complements, regardless of matrix semantic property among the residence abroad group, but only significantly to emotive-factive verbs among the non-residence abroad group.

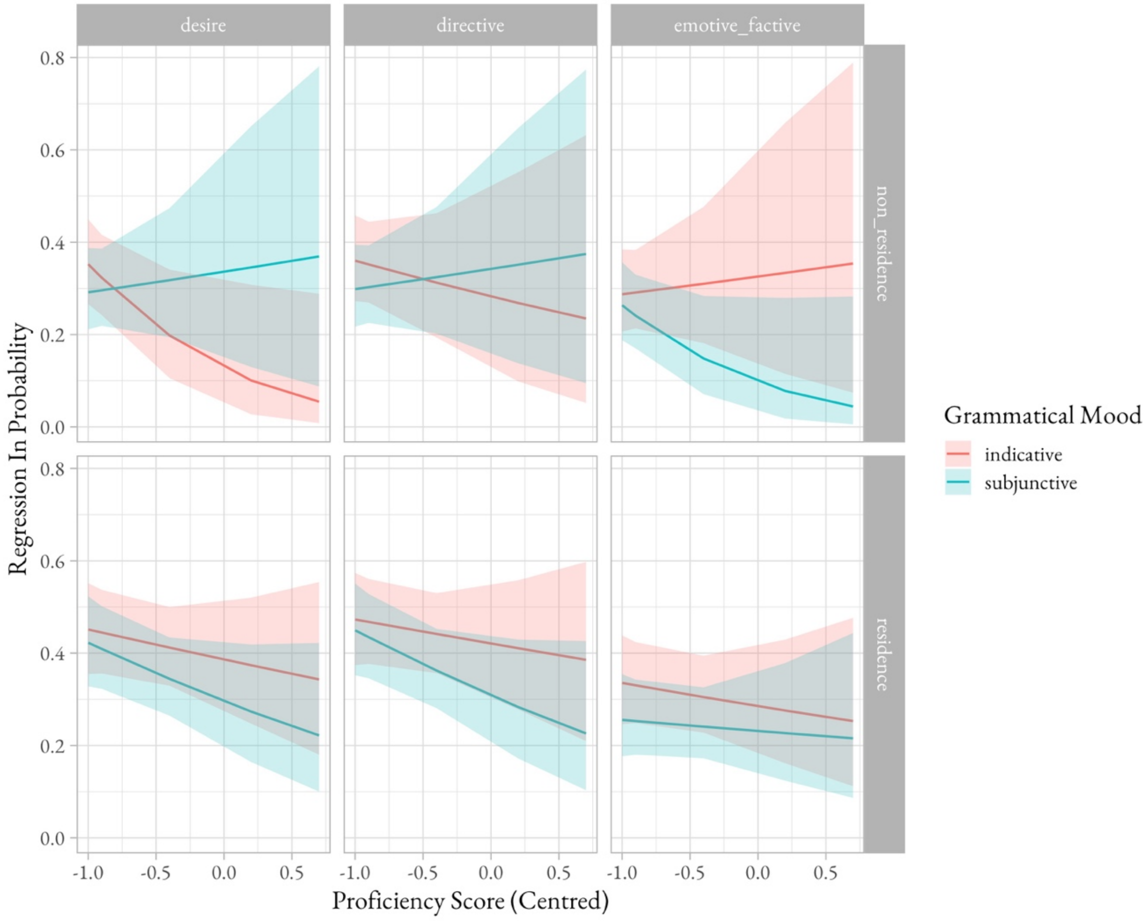


Figure 4.6 Predicted regression-in probability in the spillover region (comprehension)

Finally, for regression-out probability, there was a significant two-way interaction between condition and matrix semantic property (desire vs. directive), such that regression-out probability was higher for subjunctive (than indicative sentences) with desire verbs, but higher for indicative (than subjunctive) sentences with directive (and desire) verbs.

Table 4.20 Raw means (standard deviations) for L1 global in judgment

		Total Sentence Reading Time	Mean Fixation Count	Mean Fixation Duration	Acceptability Rating
Desire	Indicative	4210.19	13.39	241.59	2.72
		(1548.47)	(4.84)	(37.27)	(1.8)
	Subjunctive	4402.49	14.66	230.53	6.74
		(1747.21)	(5.38)	(34.05)	(0.69)
Directive	Indicative	4639.48	15.17	240.61	2.88
		(1669.98)	(5.22)	(37.13)	(1.94)
	Subjunctive	4725.38	16.17	230.69	6.77
		(1927.24)	(6.27)	(34.95)	(0.71)
Emotive-Factive	Indicative	4741.18	14.66	242.93	2.77
		(1901.26)	(5.89)	(37.86)	(1.81)
	Subjunctive	4558.44	15.26	227.39	6.73
		(2132.67)	(6.80)	(30.50)	(0.81)

Table 4.21 Raw means (standard deviations) for L2 global in judgment

		Total Sentence Reading Time	Mean Fixation Count	Mean Fixation Duration	Acceptability Rating
Desire	Indicative	6501.7	19	255.4	4.2
		(2883.5)	(8.5)	(34.6)	(2.1)
	Subjunctive	6564.1	19.5	249.2	6.3
		(3283)	(9.3)	(31.7)	(1.1)
Directive	Indicative	6938.9	20.6	255.3	4.8
		(3075.9)	(9.4)	(34)	(1.9)
	Subjunctive	7454.2	22.1	252.7	6.2
		(3432.4)	(10.1)	(35.9)	(1.3)
Emotive-Factive	Indicative	6605	19.6	250.9	4.7
		(2972.7)	(8.8)	(32.3)	(2.1)
	Subjunctive	6942.8	20.6	250.2	6.3
		(3376.3)	(9.7)	(34.8)	(1.2)

Table 4.22 Mixed-effects models for L1 global in judgment

		b	*SE*	*z*
Mean Fixation Count	(Intercept)	16.59	0.67	24.67
	Mood: Subjunctive vs. indicative	0.36	0.13	2.72
	Semantic: Desire vs. Directive	-0.33	0.36	-0.92
	Semantic: Emotive-factive vs. Directive	0.12	0.35	0.34
	Sentence Length (Centred)	1.73	0.26	6.74
	Trial Number	-0.04	0.01	-6.28
	Mood × Semantic (Desire vs. Directive)	0.16	0.19	0.84
	Mood × Semantic (Emotive-Factive vs. Directive)	-0.16	0.19	-0.83
Mean Fixation Duration	(Intercept)	5.47	0.02	304.60
	Mood: Subjunctive vs. indicative	-0.02	0.00	-6.81
	Semantic: Desire vs. Directive	0.00	0.01	-0.22
	Semantic: Emotive-factive vs. Directive	0.00	0.01	-0.36
	Sentence Length (Centred)	-0.01	0.00	-2.51
	Trial Number	0.00	0.00	-2.49
	Mood × Semantic (Desire vs. Directive)	0.00	0.00	0.36
	Mood × Semantic (Emotive-Factive vs. Directive)	-0.01	0.00	-1.18
Total Sentence Reading Time	(Intercept)	8.50	0.05	186.67
	Mood: Subjunctive vs. indicative	-0.01	0.01	-1.43
	Semantic: Desire vs. Directive	-0.02	0.02	-0.90
	Semantic: Emotive-factive vs. Directive	0.01	0.02	0.54
	Sentence Length (Centred)	0.09	0.02	5.35
	Trial Number	0.00	0.00	-8.66
	Mood × Semantic (Desire vs. Directive)	0.02	0.01	1.58
	Mood × Semantic (Emotive-Factive vs. Directive)	-0.02	0.01	-2.01
Acceptability Rating	(Intercept)	1.39	0.07	20.14
	Mood: Subjunctive vs. indicative	0.55	0.01	47.24
	Semantic: Desire vs. Directive	0.01	0.04	0.16
	Semantic: Emotive-factive vs. Directive	0.00	0.04	-0.04
	Sentence Length (Centred)	0.06	0.03	1.81
	Trial Number	0.00	0.00	-1.95
	Mood × Semantic (Desire vs. Directive)	0.01	0.02	0.71
	Mood × Semantic (Emotive-Factive vs. Directive)	0.00	0.02	-0.01

Note: Shaded cells denote statistical significance ($p < 0.05$).

Table 4.23 Mixed-effects models for L2 global in judgment

	Mean Fixation Count			Mean Fixation Duration			Total Sentence Reading Time			Acceptability Rating		
	b	*SE*	*t*	*b*	*SE*	*t*	*b*	*SE*	*t*	*b*	*SE*	*t*
(Intercept)	11.44	3.39	3.38	5.54	0.05	115.98	8.40	0.18	47.53	5.13	0.46	11.08
Mood: Subjunctive vs. Indicative	0.61	0.78	0.79	-0.03	0.01	-3.39	0.00	0.03	-0.11	1.88	0.14	13.46
Semantic: Desire vs. Directive	-0.13	0.88	-0.14	0.00	0.02	0.01	-0.06	0.05	-1.17	-0.13	0.20	-0.67
Semantic: Emotive-Factive vs. Directive	-0.40	0.88	-0.45	0.01	0.02	0.37	0.02	0.05	0.45	-0.11	0.20	-0.57
Proficiency (Centred)	-8.71	3.61	-2.41	0.00	0.05	0.03	-0.37	0.19	-1.97	-0.50	0.49	-1.01
Residence (in French-Speaking Country)	10.92	3.79	2.88	-0.04	0.05	-0.66	0.44	0.20	2.23	0.41	0.51	0.79
Sentence Length (Centred)	1.99	0.25	7.97	0.01	0.00	2.21	0.09	0.01	5.90	-0.07	0.05	-1.48
Trial Number	-0.04	0.01	-4.82	0.00	0.00	-0.61	0.00	0.00	-6.20	-0.01	0.00	-2.90
Mood × Semantic (Desire vs. Directive)	-0.15	0.85	-0.18	0.02	0.01	1.43	-0.01	0.04	-0.12	0.20	0.20	1.00
Mood × Semantic (Emotive-Factive vs. Directive)	0.23	0.85	0.27	0.00	0.01	-0.27	0.03	0.04	0.69	-0.07	0.20	-0.35
Mood × Prof.	0.41	0.83	0.49	-0.03	0.01	-2.43	-0.01	0.03	-0.43	1.00	0.15	6.74
Semantic (Desire vs. Directive) × Prof.	-0.04	0.90	-0.05	-0.01	0.02	-0.49	-0.04	0.05	-0.85	-0.05	0.21	-0.24
Semantic (Emotive-Factive vs. Directive) × Prof.	-0.67	0.90	-0.74	0.02	0.02	1.21	0.02	0.05	0.49	-0.19	0.21	-0.90
Mood × Residence	-0.48	0.88	-0.55	0.03	0.01	2.48	0.00	0.04	0.04	-0.82	0.16	-5.28
Semantic (Desire vs. Directive) × Residence	-0.80	0.95	-0.85	0.00	0.02	-0.01	0.01	0.05	0.16	-0.05	0.22	-0.21
Semantic (Emotive-Factive vs. Directive) × Residence	0.36	0.95	0.38	-0.02	0.02	-1.29	-0.01	0.05	-0.29	0.10	0.22	0.47
Proficiency × Residence	6.55	4.19	1.56	-0.03	0.06	-0.54	0.29	0.22	1.31	0.18	0.57	0.33
Mood × Semantic (Desire vs. Directive) × Prof.	0.12	0.91	0.13	0.03	0.02	1.93	0.02	0.05	0.43	0.11	0.21	0.52
Mood × Semantic (Emotive-Factive vs. Directive) × Prof.	-0.21	0.91	-0.24	-0.02	0.02	-0.97	0.00	0.05	0.00	-0.07	0.21	-0.36
Mood × Semantic (Desire vs. Directive) × Residence	0.68	0.95	0.72	-0.02	0.02	-1.20	0.01	0.05	0.21	-0.05	0.22	-0.20
Mood × Semantic (Emotive-Factive vs. Directive) × Residence	-0.79	0.95	-0.83	-0.01	0.02	-0.49	-0.06	0.05	-1.14	0.04	0.22	0.17
Mood × Prof. × Residence	-0.45	0.97	-0.46	0.02	0.01	1.94	0.01	0.04	0.31	-0.48	0.17	-2.78
Semantic (Desire vs. Directive) × Prof. × Residence	-0.55	1.05	-0.52	0.01	0.02	0.39	0.01	0.06	0.10	0.10	0.24	0.40
Semantic (Emotive-Factive vs. Directive) × Prof. × Residence	1.16	1.05	1.11	-0.04	0.02	-2.13	0.01	0.06	0.21	0.06	0.24	0.23
Mood × Semantic (Desire vs. Directive) × Prof. × Residence	1.04	1.05	0.99	-0.02	0.02	-1.07	0.02	0.06	0.37	-0.25	0.24	-1.02
Mood × Semantic (Emotive-Factive vs. Directive) × Prof. × Residence	-0.02	1.05	-0.02	-0.01	0.02	-0.35	-0.01	0.06	-0.16	0.10	0.24	0.41

Note: Shaded cells denote statistical significance ($p < 0.05$)

4.8 Results: Study 1b (Judgment)

4.8.1 Data Preparation and Analyses

Data preparation and analyses for Study 1b were identical to those of Study 1a. For a more in-depth description, see Section 4.7.1.

4.8.2 Global Measures

4.8.2.1 L1 Group

To examine the overall effect of mood and matrix semantic property, we calculated the following global measures: mean fixation duration, mean fixation count and total sentence reading time. For ease of exposition, we have also included the acceptability ratings attributed to each sentence in this section. For each measure, observed means are presented in Table 4.20 and model parameters in Table 4.22

Mean fixation count, mean fixation duration and acceptability rating all displayed a significant main effect of mood. For mean fixation count, subjunctive sentences attracted more fixations on average than indicative sentences. In contrast, indicative sentences received longer fixations on average than subjunctive sentences for mean fixation duration. Finally, the L1 speakers consistently rated subjunctive sentences as more grammatical than indicative sentences.

Although we did not find a significant main effect of mood for total sentence reading time, there was a significant two-way interaction between condition and matrix semantic property (emotive-factive vs. desire), such that reading times with directive verbs were higher for subjunctive (than indicative) sentences, whereas the reverse was true for emotive-factive verbs.

4.8.2.2 L2 Group

To examine the overall effect of mood and its interaction with matrix semantic property, proficiency and residence abroad on reading behaviour, we calculated the following global measures: mean fixation duration, mean fixation count and total sentence reading time. For ease

of exposition, we also included the acceptability ratings attributed to each sentence in this section. For each measure, observed means are presented in
Table 4.21 and model parameters in Table 4.23.

For mean fixation count and total sentence reading time, there was no significant main effect of mood. We did detect, however, a significant main effect of mood for mean fixation duration and acceptability ratings. In particular, mean fixation duration was higher for indicative (than subjunctive) sentences. This effect was further modulated by the individual fixed effects, proficiency and residence abroad but not their interaction, as evidenced by the following significant two-way interactions: Mood × Proficiency and Mood × Residence. Figure 4.7 visualises the interaction between mood and proficiency. It shows that while mean fixation duration was numerically but not significantly higher for indicative (than subjunctive) sentences, this sensitivity became more pronounced as proficiency increased. Furthermore, the distinction between indicative and subjunctive was more marked among non-resident abroad learners than residence abroad learners.

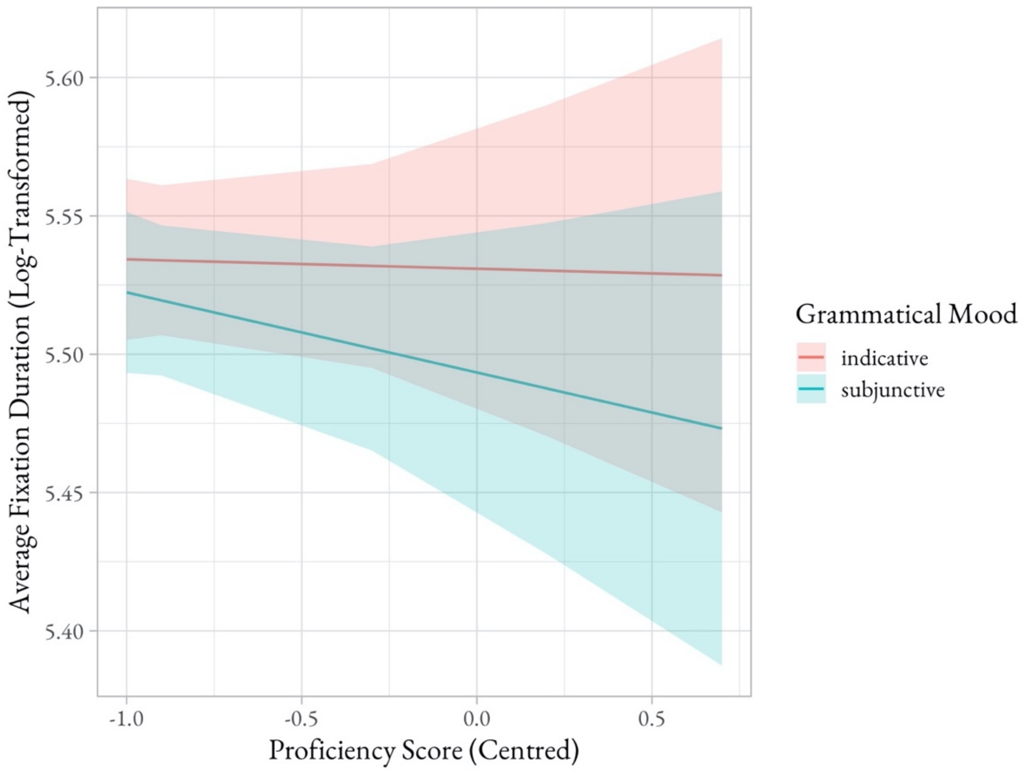


Figure 4.7 Predicted mean fixation duration (judgment)

For the acceptability ratings, there was a significant main effect of mood, such that higher ratings were attributed to subjunctive (than indicative) sentences. We also found the following significant interactions: Mood × Proficiency, Mood × Residence and Mood × Proficiency × Residence. Figure 4.8 visualises the three-way interaction. It shows that although proficiency improved L2 learners' sensitivity to mood–modality mismatches, the effect was particularly strong among non-residence abroad compared with the resident abroad learners.

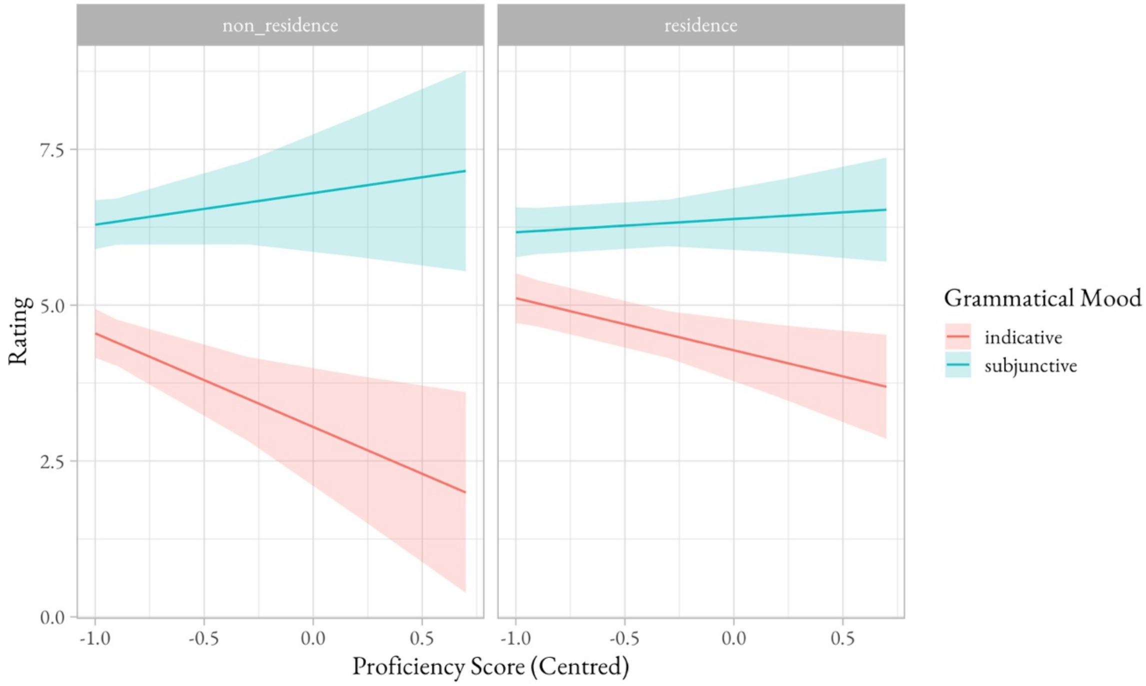


Figure 4.8 Predicted acceptability judgment ratings

Table 4.24 Raw means (standard deviations) for L1 local in judgment

		First Fixation Duration	Gaze Duration	Go-past Time	Total Reading Time	Skipping Probability	Regression-In Probability	Regression-Out Probability
				Critical Region				
Desire	Indicative	272.14 (121.51)	281.83 (126.83)	300.09 (138.38)	366.53 (186.87)	0.72 (0.45)	0.6 (0.49)	0.16 (0.36)
	Subjunctive	227.03 (76.2)	237.78 (96.24)	256.41 (121.56)	329.97 (195.02)	0.57 (0.5)	0.39 (0.49)	0.11 (0.32)
Directive	Indicative	283.15 (144.27)	296.25 (153.01)	318.39 (169.82)	465.92 (299.64)	0.64 (0.48)	0.61 (0.49)	0.12 (0.33)
	Subjunctive	243.18 (88)	256.05 (98.9)	263.12 (113.69)	330.03 (184.63)	0.46 (0.5)	0.31 (0.47)	0.06 (0.24)
Emotive-Factive	Indicative	258.95 (113.86)	292.9 (167.93)	315.27 (188.79)	436.6 (309.03)	0.67 (0.47)	0.59 (0.49)	0.1 (0.3)
	Subjunctive	238.02 (81.36)	248.7 (86.99)	255.88 (94.34)	357.58 (240.28)	0.48 (0.5)	0.27 (0.45)	0.07 (0.26)
				Spillover Region				
Desire	Indicative	263.57 (115.37)	436.93 (292.93)	557.62 (340.33)	754.04 (493.1)	0.07 (0.25)	0.26 (0.44)	0.48 (0.5)
	Subjunctive	245.28 (74.67)	414.08 (251.18)	492.31 (347.02)	732.94 (602.1)	0.07 (0.26)	0.25 (0.43)	0.23 (0.42)
Directive	Indicative	246.31 (106.52)	413.85 (322.2)	571.37 (417.03)	739.75 (526.32)	0.05 (0.22)	0.23 (0.42)	0.46 (0.5)
	Subjunctive	242.76 (86.2)	388.95 (236.61)	476.9 (318.11)	647.01 (408.5)	0.07 (0.25)	0.27 (0.44)	0.27 (0.44)
Emotive-Factive	Indicative	246.68 (99.06)	413.56 (289.65)	563.27 (392.14)	775.51 (516.31)	0.06 (0.23)	0.25 (0.44)	0.43 (0.5)
	Subjunctive	239.91 (82.36)	353.03 (215.68)	426.47 (267.07)	608.76 (434.41)	0.06 (0.24)	0.19 (0.39)	0.29 (0.45)

Table 4.25 Raw means (standard deviations) for L2 local in judgment

		First Fixation Duration	Gaze Duration	Go-past Time	Total Reading Time	Skipping Probability	Regression-In Probability	Regression-Out Probability
				Critical Region				
Desire	Indicative	262.8 (99.15)	287.33 (130.99)	316.47 (191.71)	470.12 (332.44)	0.64 (0.48)	0.48 (0.5)	0.12 (0.33)
	Subjunctive	264.49 (96.3)	281.81 (130.31)	298.95 (168.51)	452.47 (309.2)	0.46 (0.5)	0.36 (0.48)	0.14 (0.35)
Directive	Indicative	267.85 (111.09)	278.86 (115.98)	300.4 (152.2)	442.17 (327.9)	0.62 (0.49)	0.45 (0.5)	0.13 (0.33)
	Subjunctive	262.02 (106.97)	278.58 (123.17)	302.88 (143.41)	445 (266.37)	0.4 (0.49)	0.37 (0.48)	0.16 (0.37)
Emotive-Factive	Indicative	269.23 (109.6)	276.45 (118.98)	314.6 (155.09)	474.59 (326.13)	0.58 (0.5)	0.48 (0.5)	0.17 (0.37)
	Subjunctive	251.32 (92.31)	272.43 (128.66)	315.79 (186.28)	484.48 (305.28)	0.39 (0.49)	0.31 (0.46)	0.16 (0.37)
				Spillover Region				
Desire	Indicative	282.19 (107.93)	482.13 (349.68)	706.95 (556.54)	1123.71 (824.34)	0.04 (0.2)	0.41 (0.49)	0.47 (0.5)
	Subjunctive	274.35 (106.24)	516.9 (347.28)	689.75 (522.86)	1079.28 (710.94)	0.07 (0.26)	0.44 (0.5)	0.3 (0.46)
Directive	Indicative	276 (105.75)	494.36 (314.55)	658.32 (428.41)	1068.62 (767.22)	0.08 (0.27)	0.42 (0.5)	0.44 (0.5)
	Subjunctive	283.18 (113.37)	495.72 (304.44)	647.11 (432.08)	1022.48 (685.35)	0.04 (0.2)	0.43 (0.5)	0.34 (0.48)
Emotive-Factive	Indicative	263.43 (101.94)	441.87 (289.96)	601.01 (415.55)	948.76 (723.11)	0.06 (0.24)	0.34 (0.47)	0.41 (0.49)
	Subjunctive	273.43 (110.21)	480.31 (318.52)	646.25 (403.79)	933.96 (623.69)	0.07 (0.25)	0.33 (0.47)	0.42 (0.49)

Table 4.26 Mixed-effects models for L1 local in judgment

		Critical Region			Spillover Region		
		b	*SE*	*t*/*z*	*b*	*SE*	*t*/*z*
First Fixation Duration	(Intercept)	5.50	0.04	146.82	5.43	0.03	190.07
	Mood: Subjunctive vs. Indicative	-0.12	0.04	-2.85	-0.01	0.01	-0.61
	Semantic: Desire vs. Directive	-0.01	0.02	-0.44	0.03	0.02	1.99
	Semantic: Emotive-Factive vs. Directive	0.00	0.02	-0.23	-0.02	0.02	-1.00
	Region Frequency (Centred)	-0.08	0.04	-1.80	0.00	0.01	0.20
	Trial Number	0.00	0.00	-1.43	0.00	0.00	1.12
	Mood × Semantic (Desire vs. Directive)	-0.02	0.02	-1.29	-0.01	0.01	-0.88
	Mood × Semantic (Emotive-Factive vs. Directive)	0.02	0.02	0.94	0.00	0.01	0.12
Gaze Duration	(Intercept)	5.53	0.04	134.45	5.84	0.05	108.21
	Mood: Subjunctive vs. Indicative	-0.13	0.05	-2.63	-0.02	0.02	-1.03
	Semantic: Desire vs. Directive	-0.02	0.03	-0.85	0.07	0.03	2.44
	Semantic: Emotive-Factive vs. Directive	0.01	0.03	0.37	-0.06	0.03	-2.06
	Region Frequency (Centred)	-0.08	0.05	-1.61	0.16	0.02	8.58
	Trial Number	0.00	0.00	-1.17	0.00	0.00	-0.82
	Mood × Semantic (Desire vs. Directive)	-0.02	0.02	-0.87	0.01	0.02	0.24
	Mood × Semantic (Emotive-Factive vs. Directive)	0.00	0.02	0.09	-0.03	0.02	-1.34
Go-Past Time	(Intercept)	5.57	0.04	127.70	6.12	0.05	111.45
	Mood: Subjunctive vs. Indicative	-0.13	0.05	-2.45	-0.07	0.02	-4.68
	Semantic: Desire vs. Directive	-0.01	0.03	-0.39	0.04	0.04	0.90
	Semantic: Emotive-Factive vs. Directive	0.00	0.03	0.16	-0.05	0.04	-1.29
	Region Frequency (Centred)	-0.07	0.05	-1.33	0.25	0.03	8.61
	Trial Index	0.00	0.00	-0.96	0.00	0.00	-2.41
	Mood × Semantic (Desire vs. Directive)	-0.01	0.02	-0.31	0.02	0.02	0.73
	Mood × Semantic (Emotive-Factive vs. Directive)	0.00	0.02	-0.08	-0.03	0.02	-1.36
Total Reading Time	(Intercept)	5.79	0.06	94.75	6.49	0.06	102.21
	Mood: Subjunctive vs. Indicative	-0.17	0.07	-2.21	-0.06	0.02	-3.94
	Semantic: Desire vs. Directive	-0.05	0.04	-1.24	0.05	0.04	1.31
	Semantic: Emotive-Factive vs. Directive	0.03	0.04	0.75	-0.04	0.04	-1.01
	Region Frequency (Centred)	-0.09	0.08	-1.21	0.27	0.03	10.01
	Trial Index	0.00	0.00	-1.29	0.00	0.00	-4.70
	Mood × Semantic (Desire vs. Directive)	0.04	0.03	1.27	0.04	0.02	1.59
	Mood × Semantic (Emotive-Factive vs. Directive)	0.00	0.03	-0.01	-0.05	0.02	-2.34
Regression-In Probability	(Intercept)	-0.49	0.26	-1.90	-0.80	0.22	-3.58
	Mood: Subjunctive vs. Indicative	-0.51	0.24	-2.16	-0.05	0.08	-0.60
	Semantic: Desire vs. Directive	0.12	0.12	0.95	0.11	0.17	0.68
	Semantic: Emotive-Factive vs. Directive	-0.15	0.12	-1.21	-0.18	0.17	-1.04
	Region Frequency/Length (Centred)	0.13	0.24	0.53	0.12	0.12	1.03
	Trial Index	0.01	0.00	1.30	-0.01	0.00	-3.42
	Mood × Semantic (Desire vs. Directive)	0.22	0.12	1.79	0.01	0.11	0.08
	Mood × Semantic (Emotive-Factive vs. Directive)	-0.19	0.12	-1.52	-0.16	0.11	-1.41
Regression-Out Probability	(Intercept)	-2.26	0.29	-7.70	-0.77	0.25	-3.02
	Mood: Subjunctive vs. Indicative	-0.26	0.35	-0.74	-0.53	0.08	-6.94
	Semantic: Desire vs. Directive	0.34	0.17	1.97	-0.04	0.22	-0.20
	Semantic: Emotive-Factive vs. Directive	-0.20	0.19	-1.03	0.01	0.22	0.02
	Region Frequency/Length (Centred)	0.00	0.35	0.00	0.30	0.16	1.91
	Trial Index	0.00	0.01	0.11	0.00	0.00	0.24
	Mood × Semantic (Desire vs. Directive)	0.07	0.17	0.40	-0.13	0.11	-1.21
	Mood × Semantic (Emotive-Factive vs. Directive)	0.08	0.19	0.43	0.08	0.11	0.71

Note: Shaded cells denote statistical significance ($p < 0.05$).

Table 4.27 Mixed-effects models for L2 local in judgment

		Critical Region			Spillover region		
		b	*SE*	*t/z*	*b*	*SE*	*t/z*
First Fixation Duration	(Intercept)	5.45	0.07	78.22	5.55	0.06	88.01
	Mood: Subjunctive vs. Indicative	-0.12	0.05	-2.28	-0.03	0.03	-0.90
	Semantic: Desire vs. Directive	0.03	0.07	0.52	0.03	0.05	0.56
	Semantic: Emotive-Factive vs. Directive	-0.02	0.06	-0.35	0.01	0.05	0.19
	Proficiency (Centred)	-0.02	0.07	-0.31	0.01	0.06	0.23
	Residence (in French-Speaking Country)	-0.01	0.07	-0.12	-0.06	0.07	-0.92
	Region Frequency / Length (Centred)	-0.05	0.03	-1.55	0.01	0.01	1.32
	Trial Number	0.00	0.00	1.73	0.00	0.00	2.02
	Mood × Semantic (Desire vs. Directive)	-0.03	0.07	-0.42	0.05	0.05	0.95
	Mood × Semantic (Emotive-Factive vs. Directive)	0.00	0.06	-0.01	0.07	0.05	1.38
	Mood × Prof.	-0.07	0.05	-1.40	-0.06	0.04	-1.68
	Semantic (Desire vs. Directive) × Prof.	0.01	0.07	0.10	0.03	0.05	0.56
	Semantic (Emotive-Factive vs. Directive) × Prof.	-0.01	0.07	-0.10	0.04	0.05	0.85
	Mood × Residence	0.05	0.05	1.00	0.00	0.04	0.13
	Semantic (Desire vs. Directive) × Residence	-0.02	0.07	-0.33	0.01	0.05	0.13
	Semantic (Emotive-Factive vs. Directive) × Residence	0.02	0.07	0.35	-0.08	0.05	-1.47
	Proficiency × Residence	0.00	0.08	-0.05	-0.05	0.07	-0.65
	Mood × Semantic (Desire vs. Directive) × Prof.	-0.03	0.07	-0.41	0.08	0.05	1.46
	Mood × Semantic (Emotive-Factive vs. Directive) × Prof.	0.02	0.07	0.25	0.05	0.05	0.93
	Mood × Semantic (Desire vs. Directive) × Residence	0.02	0.07	0.31	-0.05	0.05	-0.96
	Mood × Semantic (Emotive-Factive vs. Directive) × Residence	-0.02	0.07	-0.29	-0.08	0.05	-1.54
	Mood × Prof. × Residence	0.02	0.05	0.39	0.04	0.04	0.91
	Semantic (Desire vs. Directive) × Prof. × Residence	0.02	0.08	0.29	0.00	0.06	-0.06
	Semantic (Emotive-Factive vs. Directive) × Prof. × Residence	-0.01	0.07	-0.11	-0.11	0.06	-1.91
	Mood × Semantic (Desire vs. Directive) × Prof. × Residence	-0.02	0.08	-0.20	-0.06	0.06	-0.93
	Mood × Semantic (Emotive-Factive vs. Directive) × Prof. × Residence	-0.02	0.07	-0.30	-0.08	0.06	-1.29
Gaze Duration	(Intercept)	5.52	0.08	71.63	5.94	0.11	53.76
	Mood: Subjunctive vs. Indicative	-0.12	0.06	-2.17	-0.10	0.05	-1.83
	Semantic: Desire vs. Directive	0.04	0.07	0.55	0.07	0.08	0.83
	Semantic: Emotive-Factive vs. Directive	-0.04	0.07	-0.67	-0.06	0.08	-0.72
	Proficiency (Centred)	0.01	0.08	0.17	-0.04	0.11	-0.34
	Residence (in French-Speaking Country)	-0.03	0.08	-0.41	-0.04	0.12	-0.36
	Region Frequency / Length (Centred)	-0.07	0.03	-2.03	0.19	0.03	7.46
	Trial Number	0.00	0.00	1.58	0.00	0.00	0.85
	Mood × Semantic (Desire vs. Directive)	0.01	0.07	0.08	-0.05	0.08	-0.61
	Mood × Semantic (Emotive-Factive vs. Directive)	-0.02	0.07	-0.27	0.18	0.08	2.31
	Mood × Prof.	-0.06	0.05	-1.13	-0.16	0.06	-2.74
	Semantic (Desire vs. Directive) × Prof.	0.01	0.08	0.13	0.06	0.08	0.68
	Semantic (Emotive-Factive vs. Directive) × Prof.	-0.03	0.07	-0.43	-0.04	0.08	-0.50
	Mood × Residence	0.03	0.05	0.59	0.10	0.06	1.56
	Semantic (Desire vs. Directive) × Residence	-0.02	0.08	-0.22	-0.04	0.09	-0.45
	Semantic (Emotive-Factive vs. Directive) × Residence	0.04	0.07	0.56	-0.06	0.09	-0.71
	Proficiency × Residence	-0.06	0.09	-0.65	-0.06	0.13	-0.44
	Mood × Semantic (Desire vs. Directive) × Prof.	-0.01	0.08	-0.11	-0.01	0.08	-0.16
	Mood × Semantic (Emotive-Factive vs. Directive) × Prof.	0.00	0.07	-0.03	0.14	0.08	1.71
	Mood × Semantic (Desire vs. Directive) × Residence	-0.04	0.08	-0.57	0.08	0.09	0.90
	Mood × Semantic (Emotive-Factive vs. Directive) × Residence	0.02	0.07	0.29	-0.19	0.09	-2.27
	Mood × Prof. × Residence	0.01	0.06	0.10	0.13	0.07	1.97
	Semantic (Desire vs. Directive) × Prof. × Residence	0.02	0.09	0.19	-0.06	0.09	-0.61
	Semantic (Emotive-Factive vs. Directive) × Prof. × Residence	0.02	0.08	0.29	0.00	0.09	-0.03
	Mood × Semantic (Desire vs. Directive) × Prof. × Residence	-0.03	0.09	-0.38	0.01	0.10	0.09
	Mood × Semantic (Emotive-Factive vs. Directive) × Prof. × Residence	0.00	0.08	-0.06	-0.10	0.09	-1.07
Go-Past Time	(Intercept)	5.59	0.09	61.32	6.32	0.13	50.10
	Mood: Subjunctive vs. Indicative	-0.10	0.06	-1.56	-0.05	0.05	-1.00
	Semantic: Desire vs. Directive	-0.04	0.08	-0.51	0.14	0.09	1.65
	Semantic: Emotive-Factive vs. Directive	0.04	0.07	0.54	-0.13	0.09	-1.52
	Proficiency (Centred)	0.00	0.09	-0.05	0.03	0.13	0.20

		Critical Region			Spillover region		
		b	*SE*	*t/z*	*b*	*SE*	*t/z*
	Residence (in French-Speaking Country)	0.01	0.10	0.13	-0.08	0.13	-0.57
	Region Frequency / Length (Centred)	-0.08	0.04	-2.27	0.25	0.03	7.18
	Trial Number	0.00	0.00	0.17	0.00	0.00	-1.64
	Mood × Semantic (Desire vs. Directive)	-0.03	0.08	-0.44	-0.10	0.08	-1.32
	Mood × Semantic (Emotive-Factive vs. Directive)	-0.03	0.07	-0.44	0.16	0.08	2.10
	Mood × Prof.	0.00	0.06	-0.01	-0.07	0.06	-1.21
	Semantic (Desire vs. Directive) × Prof.	-0.08	0.08	-0.94	0.12	0.08	1.52
	Semantic (Emotive-Factive vs. Directive) × Prof.	0.04	0.08	0.54	-0.10	0.08	-1.20
	Mood × Residence	0.00	0.06	0.06	0.05	0.06	0.87
	Semantic (Desire vs. Directive) × Residence	0.04	0.09	0.42	-0.06	0.08	-0.71
	Semantic (Emotive-Factive vs. Directive) × Residence	0.01	0.08	0.15	0.06	0.08	0.74
	Proficiency × Residence	-0.04	0.11	-0.38	-0.18	0.15	-1.23
	Mood × Semantic (Desire vs. Directive) × Prof.	-0.06	0.08	-0.75	-0.04	0.08	-0.47
	Mood × Semantic (Emotive-Factive vs. Directive) × Prof.	0.00	0.08	0.03	0.11	0.08	1.34
	Mood × Semantic (Desire vs. Directive) × Residence	-0.03	0.09	-0.40	0.09	0.09	1.03
	Mood × Semantic (Emotive-Factive vs. Directive) × Residence	0.06	0.08	0.79	-0.17	0.08	-2.01
	Mood × Prof. × Residence	-0.04	0.06	-0.60	0.05	0.07	0.73
	Semantic (Desire vs. Directive) × Prof. × Residence	0.11	0.10	1.12	-0.08	0.09	-0.80
	Semantic (Emotive-Factive vs. Directive) × Prof. × Residence	-0.03	0.09	-0.32	0.10	0.09	1.04
	Mood × Semantic (Desire vs. Directive) × Prof. × Residence	0.02	0.10	0.16	0.03	0.09	0.28
	Mood × Semantic (Emotive-Factive vs. Directive) × Prof. × Residence	0.01	0.09	0.17	-0.12	0.09	-1.31
Total Reading Time	(Intercept)	5.69	0.16	35.64	6.39	0.19	33.70
	Mood: Subjunctive vs. Indicative	-0.06	0.09	-0.69	-0.03	0.06	-0.52
	Semantic: Desire vs. Directive	0.04	0.10	0.43	0.08	0.09	0.88
	Semantic: Emotive-Factive vs. Directive	-0.02	0.09	-0.21	-0.07	0.09	-0.77
	Proficiency (Centred)	-0.14	0.17	-0.84	-0.33	0.20	-1.68
	Residence (in French-Speaking Country)	0.27	0.18	1.53	0.37	0.21	1.79
	Region Frequency / Length (Centred)	-0.03	0.06	-0.55	0.23	0.04	6.27
	Trial Number	0.00	0.00	-0.65	0.00	0.00	-3.06
	Mood × Semantic (Desire vs. Directive)	-0.09	0.10	-0.87	-0.01	0.08	-0.12
	Mood × Semantic (Emotive-Factive vs. Directive)	0.00	0.09	-0.02	0.13	0.08	1.72
	Mood × Prof.	-0.06	0.07	-0.79	-0.04	0.06	-0.64
	Semantic (Desire vs. Directive) × Prof.	0.03	0.11	0.27	-0.01	0.08	-0.12
	Semantic (Emotive-Factive vs. Directive) × Prof.	-0.06	0.10	-0.60	0.02	0.08	0.24
	Mood × Residence	0.04	0.08	0.49	0.02	0.06	0.35
	Semantic (Desire vs. Directive) × Residence	-0.06	0.11	-0.58	-0.03	0.09	-0.32
	Semantic (Emotive-Factive vs. Directive) × Residence	0.09	0.10	0.84	-0.07	0.09	-0.76
	Proficiency × Residence	0.08	0.19	0.43	0.19	0.23	0.81
	Mood × Semantic (Desire vs. Directive) × Prof.	-0.06	0.11	-0.58	0.04	0.08	0.48
	Mood × Semantic (Emotive-Factive vs. Directive) × Prof.	-0.01	0.10	-0.07	0.09	0.08	1.09
	Mood × Semantic (Desire vs. Directive) × Residence	0.09	0.11	0.83	0.09	0.09	1.02
	Mood × Semantic (Emotive-Factive vs. Directive) × Residence	-0.02	0.10	-0.20	-0.22	0.09	-2.47
	Mood × Prof. × Residence	0.02	0.08	0.28	0.04	0.07	0.54
	Semantic (Desire vs. Directive) × Prof. × Residence	-0.01	0.12	-0.11	-0.02	0.10	-0.20
	Semantic (Emotive-Factive vs. Directive) × Prof. × Residence	0.07	0.11	0.61	-0.03	0.10	-0.34
	Mood × Semantic (Desire vs. Directive) × Prof. × Residence	0.11	0.12	0.89	0.07	0.10	0.71
	Mood × Semantic (Emotive-Factive vs. Directive) × Prof. × Residence	-0.05	0.11	-0.45	-0.17	0.10	-1.78
Regression-In Probability	(Intercept)	-1.24	0.31	-4.03	-1.37	0.25	-5.48
	Mood: Subjunctive vs. Indicative	-0.07	0.33	-0.22	-0.02	0.23	-0.07
	Semantic: Desire vs. Directive	0.26	0.41	0.65	-0.52	0.33	-1.58
	Semantic: Emotive-Factive vs. Directive	-0.79	0.41	-1.95	0.12	0.32	0.36
	Proficiency (Centred)	-0.89	0.29	-3.09	-0.92	0.24	-3.91
	Residence (in French-Speaking Country)	1.00	0.31	3.26	1.15	0.25	4.61
	Region Frequency / Length (Centred)	0.36	0.17	2.05	-0.05	0.05	-0.93
	Trial Number	0.00	0.00	-1.32	0.00	0.00	-1.18
	Mood × Semantic (Desire vs. Directive)	-0.05	0.41	-0.13	0.26	0.33	0.80
	Mood × Semantic (Emotive-Factive vs. Directive)	0.04	0.41	0.11	-0.13	0.32	-0.41
	Mood × Prof.	-0.13	0.29	-0.44	-0.03	0.24	-0.12
	Semantic (Desire vs. Directive) × Prof.	0.23	0.42	0.55	-0.83	0.34	-2.44
	Semantic (Emotive-Factive vs. Directive) × Prof.	-0.69	0.41	-1.67	0.37	0.33	1.10

		Critical Region			Spillover region		
		b	*SE*	*t*/*z*	*b*	*SE*	*t*/*z*
	Mood × Residence	0.13	0.31	0.41	0.16	0.25	0.63
	Semantic (Desire vs. Directive) × Residence	-0.21	0.45	-0.47	0.46	0.36	1.26
	Semantic (Emotive-Factive vs. Directive) × Residence	0.73	0.44	1.65	-0.27	0.35	-0.77
	Proficiency × Residence	0.65	0.33	2.00	0.81	0.27	3.03
	Mood × Semantic (Desire vs. Directive) × Prof.	0.05	0.42	0.12	0.12	0.34	0.36
	Mood × Semantic (Emotive-Factive vs. Directive) × Prof.	0.10	0.41	0.25	-0.06	0.33	-0.19
	Mood × Semantic (Desire vs. Directive) × Residence	0.30	0.45	0.67	-0.06	0.36	-0.16
	Mood × Semantic (Emotive-Factive vs. Directive) × Residence	-0.36	0.44	-0.83	0.00	0.35	0.00
	Mood × Prof. × Residence	0.08	0.33	0.23	0.22	0.27	0.84
	Semantic (Desire vs. Directive) × Prof. × Residence	-0.12	0.47	-0.26	0.65	0.39	1.69
	Semantic (Emotive-Factive vs. Directive) × Prof. × Residence	0.66	0.46	1.43	-0.16	0.38	-0.43
	Mood × Semantic (Desire vs. Directive) × Prof. × Residence	0.26	0.47	0.55	0.25	0.39	0.66
	Mood × Semantic (Emotive-Factive vs. Directive) × Prof. × Residence	-0.42	0.46	-0.92	-0.13	0.38	-0.33
Regression-Out	(Intercept)	-1.59	0.41	-3.83	-0.09	0.23	-0.40
Probability	Mood: Subjunctive vs. Indicative	-0.35	0.44	-0.79	-0.58	0.21	-2.79
	Semantic: Desire vs. Directive	-0.76	0.60	-1.28	-0.37	0.30	-1.22
	Semantic: Emotive-Factive vs. Directive	0.47	0.49	0.96	0.17	0.29	0.59
	Proficiency (Centred)	0.11	0.40	0.29	0.48	0.22	2.16
	Residence (in French-Speaking Country)	-0.36	0.43	-0.84	-0.45	0.23	-1.93
	Region Frequency / Length (Centred)	-0.29	0.24	-1.21	0.13	0.05	2.54
	Trial Number	-0.01	0.00	-1.40	0.00	0.00	0.37
	Mood × Semantic (Desire vs. Directive)	-0.60	0.60	-1.01	-0.52	0.30	-1.72
	Mood × Semantic (Emotive-Factive vs. Directive)	-0.43	0.49	-0.87	0.16	0.29	0.57
	Mood × Prof.	0.18	0.40	0.44	-0.39	0.22	-1.73
	Semantic (Desire vs. Directive) × Prof.	-0.74	0.60	-1.23	-0.28	0.32	-0.87
	Semantic (Emotive-Factive vs. Directive) × Prof.	0.31	0.53	0.60	0.12	0.31	0.38
	Mood × Residence	0.31	0.43	0.72	0.43	0.23	1.85
	Semantic (Desire vs. Directive) × Residence	0.39	0.68	0.57	0.49	0.34	1.46
	Semantic (Emotive-Factive vs. Directive) × Residence	-0.06	0.56	-0.11	-0.09	0.32	-0.26
	Proficiency × Residence	-0.79	0.46	-1.71	-0.63	0.26	-2.45
	Mood × Semantic (Desire vs. Directive) × Prof.	-0.69	0.60	-1.14	-0.43	0.32	-1.34
	Mood × Semantic (Emotive-Factive vs. Directive) × Prof.	-0.39	0.53	-0.73	0.01	0.31	0.02
	Mood × Semantic (Desire vs. Directive) × Residence	0.42	0.68	0.62	0.39	0.34	1.17
	Mood × Semantic (Emotive-Factive vs. Directive) × Residence	0.47	0.56	0.83	-0.08	0.32	-0.24
	Mood × Prof. × Residence	-0.18	0.46	-0.40	0.36	0.26	1.39
	Semantic (Desire vs. Directive) × Prof. × Residence	0.36	0.70	0.52	0.49	0.37	1.31
	Semantic (Emotive-Factive vs. Directive) × Prof. × Residence	0.01	0.61	0.02	-0.12	0.36	-0.32
	Mood × Semantic (Desire vs. Directive) × Prof. × Residence	0.43	0.70	0.62	0.54	0.37	1.46
	Mood × Semantic (Emotive-Factive vs. Directive) × Prof. × Residence	0.51	0.61	0.83	-0.19	0.36	-0.54

Note: Shaded cells denote statistical significance (p < .05).

4.8.3 Local Measures

4.8.3.1 L1 Group

To examine the overall effect of mood and its interaction with matrix semantic property on reading behaviour, we calculated the following local measures: first fixation duration, gaze duration, go-past time, total reading time, regression-in probability and regression-out probability. For a more in-depth description of these measures, see Section 2.2.4.1. The means (and standard deviations) and the fixed effect estimates for all the local measures in the critical and spillover region are shown in Table 4.24 and Table 4.26, respectively.

In the critical region, we found a significant main effect of mood for first fixation duration, gaze duration, go-past time, total reading time and regression-in probability. These reading measures were higher for indicative (than subjunctive) sentences.

In the spillover region, there was a significant main effect of mood for go-past time, total reading time and regression-out probability. These reading measures were higher for indicative than subjunctive sentences. For total reading time, we detected a significant interaction between mood and matrix semantic property (emotive-factive vs. directive), such that the distinction between the indicative and subjunctive was stronger among emotive-factive than directive verbs.

4.8.3.2 L2 Group

To examine L2 learners' sensitivity to mood and the influence of matrix semantic property, proficiency and residence abroad on reading behaviour, we calculated the following local measures: first fixation duration, gaze duration, go-past time, total reading time, regression-in probability and regression-out probability. For a more in-depth description of these measures, see Section 2.2.4.1. The means (and standard deviations) and the fixed effect estimates for all the local measures in the critical and spillover region are shown in Table 4.25 and Table 4.27, respectively.

In the critical region, we found a significant main effect of mood for first fixation duration and gaze duration, such that reading times were higher for indicative than subjunctive sentences. The other reading measures did not exhibit a significant main effect of mood.

In the spillover region, there were two significant interactions for gaze duration: Mood × Semantic (Emotive-Factive vs. Directive) and Mood × Proficiency. In particular, gaze duration was longer for subjunctive than indicative sentences with emotive-factive and desire verbs, but longer for indicative than subjunctive sentences with directive verbs. Figure 4.9 visualises the interaction between mood and proficiency. It shows that at lower proficiency levels, L2 learners exhibited longer gaze durations for subjunctive than indicative sentences, but as proficiency increased, this pattern reversed to the extent that gaze duration was numerically but not significantly longer for indicative than subjunctive sentences.

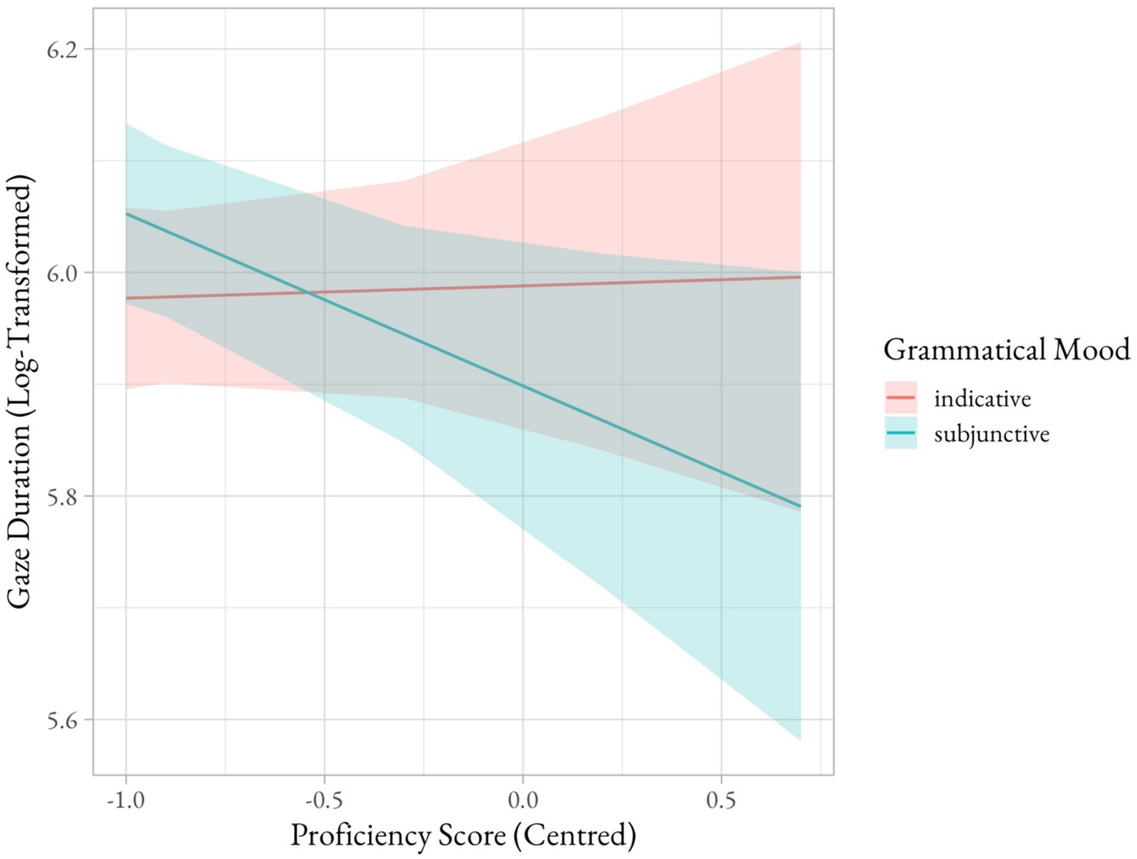


Figure 4.9 Predicted gaze duration in the spillover region (judgment)

For both go-past time and total reading time, we observed a significant three-way interaction between mood, matrix semantic property (emotive-factive vs. directive) and residence abroad. Among the non-residence abroad group, go-past time and total reading time were longer for indicative than subjunctive sentences with desire and directive verbs, but not emotive-factive verbs. Among the residence abroad group, however, go-past times were consistently longer for subjunctive than indicative verbs, regardless of matrix semantic property. In contrast, total reading time was longer for indicative than subjunctive sentences with desire and emotive-factive verbs, but not directive verbs.

4.9 Discussion

The data presented in Sections 4.7 and 4.8 revealed several key findings. First, native speakers of French demonstrated both offline and online sensitivity to mood–modality mismatches, regardless of the lexical-semantic properties of the matrix verb. Second, L2 learners of French exhibited offline and some online sensitivity to mood–modality mismatches. In particular, these learners were more likely to display online sensitivity to mood–modality mismatches with directive and desire verbs than emotive-factive verbs. This effect was further modulated by proficiency and residence abroad, suggesting that online sensitivity among L2 learners depended, in large part, on the L1 properties, proficiency and residence abroad. The exact nature of this behaviour will be discussed in the current section.

EXPERIMENTAL PREDICTIONS

Recall the initial hypotheses concerning the underlying cause of differences between the native speakers and the second language learners of French, reformulated below:

- Hypothesis 1: L2 learners should be more sensitive to mood–modality mismatches in the offline than the online data.
- Hypothesis 2: If sensitivity to mood–modality mismatches is modulated by the L1 properties, then L2 learners should demonstrate sensitivity first with directive predicates, then desire predicates and finally emotive-factive predicates.
- Hypothesis 3: As proficiency increases, L2 learners should exhibit sensitivity to mood–modality mismatches not only with directive predicates, but also desire and emotive-factive predicates.
- Hypothesis 4: Exposure to (naturalistic) input, through residence abroad, should improve sensitivity to mood–modality mismatches.

The findings from this study attest to an asymmetry between offline and online performance among the L2 group, offering support for Hypothesis 1. In particular, we found that the L2 learners were sensitive to mood–modality mismatches in the offline data, regardless of matrix semantic property, proficiency and/or residence abroad. Although proficiency did not exclusively determine sensitivity to mood–modality mismatches, it did modulate the magnitude of the effect, such that higher proficiency yielded increased sensitivity. This finding supports Hypothesis 3 and is in line with previous studies showing that subjunctive knowledge improves with proficiency (Howard, 2008; Bartning, Lundell and Hancock, 2012; Ayoun, 2013; McManus and Mitchell, 2015). Furthermore, while both residence and non-residence abroad learners were sensitive to mood–modality mismatches, the effect was stronger among non-residence abroad learners. This finding contradicts Hypothesis 4 and will be explored in more detail in the latter part of this chapter.

Sensitivity, however, was limited and highly constrained in the online data, as we will discuss. This asymmetry between offline and online performance is consistent with previous studies showing that despite target-like offline knowledge, L2 speakers may still exhibit non-target-like performance during real-time processing (e.g., Papadopoulou and Clahsen, 2003; Marinis *et al.*, 2005; Shimanskaya and Slabakova, 2017; Stepanov *et al.*, 2019). Furthermore, this indicates a dissociation between processing patterns and grammatical knowledge comparable to that observed in similar studies on immersed and highly proficient learners (Harrington, 2001; Juffs, 2001).

More generally, it could be argued that this asymmetry between offline and online behaviour suggests that the grammar and language processing constitute two independent systems. However, as Lewis and Phillips (2015) state, a one-system hypothesis can easily accommodate the dissociation between parsing performance and grammatical knowledge. For example, it is possible that the real-time mechanisms sub-serving the grammar may not be adequately equipped to handle the task in question, particularly when such a task requires domain-general resources, such as working memory and cognitive control (Stepanov *et al.*, 2019). This means that an asymmetry between offline and online performance does not necessarily reflect an absence of grammatical knowledge *per se*, but rather difficulties applying said grammatical knowledge during real-time processing due to limited domain-general resources (e.g., McDonald, 2006; Hopp, 2010; Sorace, 2011; Dekydtspotter and Renaud, 2014).

As stated above, online sensitivity among L2 learners was highly constrained by several factors, including matrix semantic property, proficiency and residence abroad. It is important to highlight that it was often the combination of these factors, rather than any one factor alone, that predicted online sensitivity to mood–modality mismatches among L2 learners. In fact, it was only in the judgment task that we detected three online measures (mean fixation duration and first fixation duration and gaze duration in the critical region) where L2 learners exhibited sensitivity to mood–modality mismatches, irrespective of the matrix semantic property, proficiency or residence abroad. This finding suggests that when reading for judgment, L2 speakers were sensitive to the experimental manipulation not only in the early stages of processing, but also on a more global level. In particular, L2 speakers were able to demonstrate target-like sensitivity to both the lexical properties (as evidenced by first fixation duration) and the morphosyntactic violation (as shown by gaze duration) in the critical region, regardless of the matrix semantic property, proficiency or residence abroad. Study 3 will further explore why sensitivity effects were only observed in the judgment task, and not the comprehension task.

Before discussing the influence of the L1 and its interaction with additional factors, such as proficiency and residence abroad, on L2 sensitivity to mood–modality mismatches, we must first examine the effect of proficiency itself. As we have previously mentioned, proficiency predicted the strength of L2 sensitivity to mood–modality mismatches in the offline data (that is, grammaticality judgment ratings). A similar effect was observed for mean fixation duration in the judgment task.

While proficiency modulated the strength of sensitivity for the grammaticality judgment ratings and mean fixation duration, it in fact determined sensitivity for gaze duration in the spillover region of the judgment task. Such a finding suggests that while proficiency improves L2 learners' ability to detect morphosyntactic anomalies (as evidenced through gaze duration), sensitivity is both task dependent and often delayed, only surfacing in the spillover region when learners are instructed to focus on form rather than content. The influence of task design will be further explored in Study 3. More generally, these findings indicate that higher proficiency yielded increased sensitivity to morphosyntactic violations. This supports Hypothesis 3 and is in line with previous studies showing that proficiency significantly improves the efficiency and accuracy of L2 processing (McDonald, 2000, 2006; Hopp, 2006; Jackson, 2008; Foote, 2011; Lim and Christianson, 2015).

As we briefly mentioned in the introduction, a key finding from the current study was the influence of the lexical-semantic properties of the matrix verb, and by extension the L1, on L2 sensitivity to mood–modality mismatches. In the few instances where we did find sensitivity to mood–modality mismatches, it was predominately with directive and desire predicates. This apparent L1 effect, however, was highly constrained in itself by proficiency and residence abroad and did not surface in every measure, whether on a local or a global level.

In fact, we only observed one measure (regression-out probability in the spillover region of the comprehension task) where the L1 properties alone predicted L2 sensitivity to mood–modality mismatches. Here, regressions were more likely to be made from the spillover region to previous parts of the sentence when processing indicative (than subjunctive) complements to directive predicates[23], but not with emotive-factive predicates in the comprehension task.

[23] And also desire predicates. However, our mixed-effects models did not compare desire and emotive-factive verbs.

Given that regression-out probability has been shown to index (late) processing difficulty (Pearlmutter, Garnsey and Bock, 1999), such a finding reflects an emerging ability among L2 learners to undertake reanalysis (in order to identify the source of the ungrammaticality) following integration difficulties, an ability that seems to be almost exclusively restricted to directive verbs. Assuming a one-system hypothesis, this finding suggests that these L2 learners have access, at least partially, to the relevant grammatical features required to parse mood–modality mismatches with directive and desire predicates. However, as we will discuss, the influence of the L1 was often found in combination with other factors, which indicates that the L1 properties alone cannot explain sensitivity to mood–modality mismatches.

Turning now to the interaction between the matrix semantic properties and proficiency, the current dataset suggests that these two factors in fact predicted sensitivity to mood–modality mismatches for mean fixation duration in the comprehension task. At the lowest proficiency levels tested, learners did not demonstrate sensitivity to mood–modality mismatches with directive predicates. Sensitivity did emerge, however, as a function of proficiency. This contrasted with emotive-factive predicates, for which learners did not exhibit sensitivity to mood–modality mismatches, regardless of proficiency.[24]

This finding supports the set of theories, such as the 'Full Transfer/Full Access/Full Parse' Model (Schwartz and Sprouse, 1996; White, 2000; Dekydtspotter, Schwartz and Sprouse, 2006), and the Fundamental Identity Hypothesis (Hopp, 2007), both of which assume that the L1 and L2 processing systems are fundamentally the same, regardless of differences in age, proficiency and other such factors. These models predict that as proficiency increases, L2 processing patterns will approximate, or in some cases align with, L1 processing patterns since both L1 and L2 processing are fundamentally sub-served by the same neurocognitive systems. Under such a theory, L1 transfer effects and performance factors (e.g., limited computational resources) can account for L1/L2 processing differences (e.g., McDonald, 2006; Hopp, 2010; Sorace, 2011; Dekydtspotter and Renaud, 2014). The exact nature of these transfer effects will be discussed in more detail later in the discussion.

A particularly striking finding from this study was the contradictory role that residence abroad played in predicting sensitivity to mood–modality mismatches. As we have previously discussed, sensitivity to mood–modality mismatches in the offline data was stronger among the

[24] L2 learners did not demonstrate any sensitivity with desire predicates, regardless of proficiency. Note, however, that we did not find any significant differences between desire and directive predicates in terms of L2 sensitivity to mood–modality mismatches for mean fixation duration.

non-residence abroad group than the residence abroad group, which contradicted our initial predictions (see Hypothesis 4). In the online data, we found not only evidence to suggest a facilitative, but also an (indirect) non-facilitative influence of residence abroad. In fact, in some cases, these contradictory effects were detected in the same reading measure. This evidence will be discussed in more depth below.

Interestingly, evidence in support of the facilitative effect of residence abroad was often observed in its interaction with the lexical-semantic properties of the matrix predicate and/or proficiency. For example, with mean fixation duration in the comprehension task, L2 learners demonstrated sensitivity to mood–modality mismatches with directive predicates, regardless of residence abroad experience, whereas we only found evidence of sensitivity with emotive-factive verbs in the residence abroad group.[25]

A similar finding was detected for regression-in probability[26] in the spillover region of the comprehension task. In particular, we saw that the residence abroad learners showed sensitivity to mood-modality mismatches, regardless of matrix semantic property, whereas the non-residence abroad learners were only sensitive to mood-modality mismatches with emotive-factive verbs at higher levels of proficiency. Given that regression-in probability typically indexes later processing (Inhoff, 1984; Duffy, Morris and Rayner, 1988; Rayner, 1998), such a finding reflects an (emerging, in the case of non-residence abroad learners) ability to perform reanalysis following integration difficulties. Likewise, for total reading time in the spillover region of the comprehension task, L2 learners demonstrated sensitivity to mood–modality mismatches with directive predicates, irrespective of residence abroad experience. While non-residence abroad learners did not exhibit sensitivity with emotive-factive predicates, regardless of proficiency, highly proficient residence abroad learners did. Given the lack of significant effects in the early measures for the spillover region, this suggests a late effect on processing (Liversedge, Paterson and Pickering, 1998).

Although the data did, to a certain extent, attest to a facilitative effect of residence abroad on sensitivity to mood–modality mismatches, there was evidence to suggest an indirect non-

[25] It is interesting that our analysis did not reveal a significant difference between desire and directive verbs, even though mean fixation duration was longer for subjunctive (than indicative) sentences with desire verbs, but longer for indicative (than subjunctive) sentences with directive verbs.

[26] In this study, if participants demonstrated sensitivity to mood–modality mismatches with regression-in probability, we expected that participants would be more likely to regress back to the region in question from later parts of the sentences with indicative than subjunctive sentences.

facilitative role,[27] such that non-residence abroad learners were more sensitive than residence abroad learners in some measures. Although the L2 group generally demonstrated sensitivity in mean fixation duration and acceptability ratings in the judgment task, this effect was stronger among the non-residence abroad group than the residence-abroad group.

It was typically in its interaction with other factors, such as the lexical-semantic properties and/or proficiency, however, that the non-facilitative effect of residence abroad was observed. For example, go-past times were only sensitive to mood–modality mismatches in the spillover region of both the comprehension and judgment task for the non-residence abroad group with desire and directive verbs.[28] In the comprehension task, but not the judgment task, this behaviour was modulated by proficiency, such that higher proficiency levels yielded increased sensitivity. This sensitivity found in go-past time is likely to reflect higher order processing, such as reanalysis following integration difficulties (Rayner, 1998; Rayner and Pollatsek, 2006).

In some measures, however, residence abroad appeared to have both a facilitative and a non-facilitative effect on sensitivity depending on the lexical-semantic properties of the matrix predicate. For example, the non-residence abroad group, but not the residence abroad group, showed sensitivity to mood–modality mismatches with directive predicates for total reading time in the spillover region of the judgment task.[29] In contrast, it was the residence abroad group, and not the non-residence abroad group, that demonstrated sensitivity with emotive-factive predicates. Such a finding suggests that residence abroad had an indirect non-facilitative effect on the ability to perform reanalysis (following integration difficulties) with directive verbs, but a facilitative effect with emotive-factive verbs. This finding is anomalous in light of the previously presented evidence showing that sensitivity was more likely to surface with directive (and in most cases desire verbs) than emotive-factive verbs.

As stated in the introduction, a secondary aim of this study was to explore the extent to which residence abroad, and thus naturalistic exposure to the target language, modulated L2 sensitivity to mood–modality mismatches. To our knowledge, little research has investigated how residence abroad might impact on processing patterns among L2 learners, with the exception of LaBrozzi (2012), for example. This consequently makes it somewhat challenging

[27] For ease of exposition, we refer to this as an indirect non-facilitative role due to the cross-sectional nature of the dataset. It is not the case that residence abroad weakened the performance of the residence abroad learners, but rather that we found between-group differences.

[28] A similar pattern was detected with desire predicates.

[29] Total reading time was higher for indicative than subjunctive sentences with desire verbs, regardless of residence abroad. However, we did not find a significant difference between desire and directive verbs.

to interpret the contradictory role of residence abroad and naturalistic exposure, more generally. Nevertheless, our findings largely support previous studies showing that residence abroad offers limited benefits to L2 learners and that classroom instruction generally has the more beneficial impact (Collentine, 2004; Isabelli-García, 2010; Freed, Lazar and So, 2013; Arnett, 2013). It could be argued that residence abroad interferes with the complex configurations of the grammar since higher knowledge may be noticed but not necessarily analysed accurately as such and thus not internalised into the grammar (Dekydtspotter and Renaud, 2014).

Based on the discussion above, we were unable to determine with any certainty whether residence abroad had a facilitative or a non-facilitative effect on L2 sensitivity during real-time processing. The cross-sectional nature of our dataset prevented us from exploring whether this inconclusive effect was due to the naturalistic exposure itself, or rather the profile of the L2 learners tested in this study. By this, we refer to the multifaceted nature of both the residence abroad experience and profile of L2 learners. Previous research has shown that residence abroad experiences can vary greatly between learners, in terms of the quality and quantity of exposure each learner receives (Mitchell, Tracy-Ventura and Mcmanus, 2017). Likewise, L2 learners, regardless of residence in a French-speaking country, by their very nature, exhibit individual differences in their exposure to the L2. For example, as the Language History Questionnaire revealed, some learners were more likely to interact with, and by extension, expose themselves to the target language than others. Unfortunately, due to the sample size of the current study, we were unable to analyse how the multifaceted nature of L2 exposure influenced development, using the information from the LHQ. Furthermore, it is worth noting that in recent years, an increasing number of studies have investigated the extent to which individual differences (e.g., working memory and cognitive control) lead to differences in processing patterns and ultimate attainment in L2 acquisition (Perfetti, 2007; Traxler and Tooley, 2007; Dyke, Johns and Kukona, 2014; Hopp, 2017b).

Returning now to the influence of the L1 on sensitivity to mood–modality mismatches, we found that although L2 learners only demonstrated limited sensitivity to mood–modality mismatches, this sensitivity was highly constrained by multiple factors, including the matrix semantic properties and proficiency. In particular, L2 learners were more likely to demonstrate sensitivity with directive and desire verbs, but not (so much) with emotive-factive verbs.[30]

[30] Interestingly, our analysis did not reveal a significant difference between directive and desire verbs in terms of sensitivity to mood–modality mismatches with any of the global or local measures.

Such a finding confirms our initial hypothesis that the L1 form-to-meaning mappings would influence L2 acquisition and processing. In particular, we anticipated that out of the three lexical-semantic categories, directive predicates would be the least challenging for L2 learners, since successful acquisition simply requires feature remapping. These L2 learners already have access to the functional projection, MoodP, in the mental lexicon via directive predicates in British English which allow both indicative- and subjunctive CPs as complements. As such, L2 speakers were required to move from an L1 superset relationship with a many-to-one form-to-meaning mapping to a L2 subset with a one-to-one form-to meaning mapping. Put simply, where the L1 allows both indicative and subjunctive clauses as complements to directive verbs, the L2 only permits subjunctive complements. The data from the current study confirms this prediction.

In contrast, emotive-factive predicates were predicted to be the most challenging of the three lexical-semantic categories for L2 learners. As we have discussed, L2 learners already have access to the functional projection, MoodP, and the associated features in the L1 via directive predicates. However, unlike the L2 French, these syntactic features are not mapped onto emotive-factive predicates in the L1 (British English). L2 speakers must therefore acquire new licensing conditions, by remapping and reconfiguring the existing L1 feature bundles from indicative CPs without the relevant modal feature onto a new morphosyntactic form (subjunctive CPs) that carries the modal feature, [M]. As we have seen in this discussion, the current study provides evidence to support the claim that the emotive-factive condition would be the most challenging.

The facilitative role of the L1 can also explain the (mostly) target-like sensitivity with desire verbs. Recall that in Section 3.2.1, we discussed how *for-to* infinitives can be analysed as alternative markers of modality. In particular, it was argued that complement clauses to desire predicates already mark modality in the L1 via the modal feature attached to the preposition, *for*, but that L2 learners must remap this modal feature onto another marker of modality using overt surface morphology, that is, subjunctive CPs. This meant that successful acquisition was not as challenging as it was in the case of emotive-factive predicates, but also not as easy as it was in the case of directive predicates.

Although these findings largely appear to support the Feature Reassembly framework and are consistent with recent studies by Guijarro-Fuentes (2012), Gil and Marsden (2013), Hwang and Lardiere (2013), Cho and Slabakova (2014), Domínguez, Arche and Myles (2017), among others, it is important to discuss alternative explanations. One explanation for the sensitivity effects among desire predicates may in part relate to the frequency of the matrix predicate in the L2 input. Research by Slabakova (2015) and Hopp *et al.* (2020) has found that construction frequency can determine convergence with the target grammar, to the extent that lower construction frequency in the input is more likely to result in non-target-like behaviour. For example, Section 4.2.1 revealed that desire verbs were among the most frequently occurring subjunctive triggers in the input, suggesting that the activation threshold levels of these predicates and their related feature bundles were higher than lower frequency predicates. While the frequency of the matrix predicate may explain the L2 processing patterns with desire predicates, it fails to account for directive predicates, some of which were among the least frequently occurring predicates. Unfortunately, the current study was not set up to test precisely this and as such this account is only speculative at best.

Finally, it is worth drawing attention to the location of sensitivity effects detected in this study since most of the evidence was predominately concentrated in the spillover region. In fact, the only evidence of sensitivity in the critical region we found was the two examples discussed earlier, both of which were observed in the judgment task. This suggests that sensitivity to mood–modality mismatches, and morphosyntactic anomaly detection, more generally, is typically delayed among L2 learners. This finding is in line with previous studies showing delayed sensitivity effects among L2 learners (Felser and Cunnings, 2012; Felser *et al.*, 2012; Lim and Christianson, 2015; Boxell and Felser, 2017). It could even be argued that such a finding is indicative of a latency effect arising from quantitative L1/L2 differences in terms of processing speed (i.e., slower and less automatic) or capacity (cf., McDonald, 2000, 2006; Hopp, 2007).

In conclusion, the current study revealed a noticeable asymmetry between offline and online sensitivity to mood–modality mismatches. While L2 learners consistently demonstrated offline knowledge of the subjunctive, this was not always the case for online knowledge. We found that online sensitivity was highly constrained by several factors, including L1 influence, proficiency and residence abroad. In particular, target-like sensitivity was more likely to be observed with desire and directive predicates than emotive-factive predicates. We argue that these findings are largely consistent with a processing-based extension of the FRH and show that the acquisition of structures involving feature remapping and reassembly are more likely to be a source of difficulty for L2 learners, but that these difficulties can eventually be overcome as proficiency increases. The role of residence abroad, however, was inconclusive. In summary, these findings have shown that the Feature Reassembly framework generates testable hypotheses that can enable us to develop a greater understanding of the problematic nature of particular grammatical properties.

Chapter 5 Study 2: Polarity Subjunctives

5.1 Introduction

In the previous chapter, we investigated how differences in L1/L2 feature composition impact on offline and online sensitivity to mood–modality mismatches, focusing in particular on morphosyntactic and semantic features. The current chapter extends the scope of the previous chapter by looking at the polarity subjunctive, a structure that crucially depends on a discourse-based feature.

As discussed in earlier chapters, research has consistently shown that linguistic structures involving discourse-based constraints can be a source of difficulty for many L2 learners across a range of proficiencies. The Interface Hypothesis, for example, proposes that L2 learners may experience difficulties with linguistic properties that lie at the interface between syntactic and other cognitive domains (namely, discourse-pragmatics). Sorace (2003, 2004, 2005, 2011, 2012, 2014) and colleagues (Sorace and Filiaci, 2006; Tsimpli and Sorace, 2006; Belletti, Bennati and Sorace, 2007; Sorace and Serratrice, 2009; Filiaci, Sorace and Carreiras, 2014), who adopt a modular view of language, argue for a dichotomous relation between internal and external interfaces with respect to the extent of fossilisation and non-target-like behaviour (White, 2009, 2011). In their view, external interfaces (i.e., the interface between syntax and other cognitive domains such as discourse and/or pragmatics) are more problematic than internal interfaces (i.e., the interface between syntax and other sub-modules of language such as semantics, phonetics/phonology and/or morphology). It has been argued that these difficulties result from the processing demands that interface conditions place on the parser due to insufficient computational resources. Despite such explanations, few studies have tested the predictions of this hypothesis using online methodologies, such as self-paced reading and eye-tracking.

The Feature Reassembly Hypothesis (Lardiere, 2009), however, does not predict difficulties based on the implication of interface conditions. This hypothesis stipulates that L2 learners begin language learning with "an entrenched system of morphosyntactic features already assembled into lexical items" (Hwang and Lardiere, 2013, p. 58). In order to acquire L2-specific features successfully, learners must go through two stages. First, they must perform a contrastive analysis between L1 and L2 properties in an attempt to map, either accurately or inaccurately, the L1 form(s) onto the L2 equivalent(s). If L1-L2 form-to-meaning mapping mismatches occur, L2 learners must then reassemble existing L1 feature configurations into new L2 feature bundles and lexical items (Hwang and Lardiere, 2013, p. 58). Linguistic properties with divergent L1/L2 feature configurations are expected to be problematic given the need for feature reassembly.

The polarity subjunctive represents an interesting test case for these hypotheses, since it requires speakers to integrate discourse contextual information during real-time processing. As we have discussed in the previous chapter, most studies to date have focused on the subjunctive in obligatory contexts with few investigating contexts where the indicative can alternate with the subjunctive without ungrammaticality. The current chapter presents findings from Study 2, which sought to address the following research questions:

- Do English-speaking L2 learners of French exhibit offline and online sensitivity to the syntactic and discourse/pragmatic constraints of the subjunctive in polarity contexts?
- To what extent is this sensitivity influenced by L1/L2 differences in
    - proficiency and
    - residence in a French-speaking country?

The current chapter begins by describing the linguistic properties of the polarity subjunctive in French and English (Section 5.2.1) as well as research conducted on the L2 acquisition of the polarity subjunctive (Section 5.2.2) and on the processing of counterfactual thought and negation among monolinguals (Section 5.2.3). The research questions and predictions are then laid out in Section 5.3, with the experimental design and methodology presented in Section 5.4. Finally, the results are described in Section 5.5 (comprehension task) and Section 5.6 (judgment task), with a discussion of the implications in Section 5.7

5.2 Polarity Subjunctive in French and English

In Study 1, we discussed how the lexical-semantic properties of the matrix predicate determines subjunctive use in obligatory contexts. In polarity (broadly defined as the opposition between positive and negative forms) contexts, however, the subjunctive is licensed by a syntactic operation, such as interrogation or negation (Stowell, 1993). In French, it is restricted to epistemic verbs. Typically, these verbs require the indicative in affirmative clauses, as in (28). However, when these predicates occur in the interrogative or negated form, the subjunctive can be used, as illustrated in (29) and (30).

(28) Il croit que Pierre est / *soit revenu.

('He believes that Pierre returned$_{\text{-IND}}$ / returned$_{\text{-SUBJ}}$')

(29) Croit-il que Pierre est / soit revenu?

('Does he believe that Pierre returned$_{\text{-IND}}$ / returned$_{\text{-SUBJ}}$')

(30) Il ne croit pas que Pierre est / soit revenu.

('He does not believe that Pierre returned$_{\text{-IND}}$ / returned$_{\text{-SUBJ}}$')

Examples from Abouda (2002, p. 1) and McManus and Mitchell (2015, p. 47)

According to Quer (1998, 2001), the use of the subjunctive with epistemics in polarity contexts results from a shift in the model of evaluation with respect to the individual anchor, that is, from the speaker's epistemic model to the matrix subject's epistemic model. In other words, mood choice allows the speaker to approach the event described in the embedded clause from their own viewpoint or from the viewpoint of the matrix subject. Such an analysis therefore means that the indicative can be used in complement clauses following negated and interrogative epistemics. Here, the evaluation model would continue to be the speaker's epistemic model and the indicative would express a speaker's commitment to the truth-value of the complement clause, whereas the subjunctive would suggest a lack of such commitment.

With respect to the syntactic structure of polarity subjunctives, the modal feature in Force is valued as interpretable and the interaction between the modal feature in Force and the Mood head in the inflectional field is that of identification (Iverson, Kempchinsky and Rothman, 2008; Kempchinsky, 2009). Such an analysis can be illustrated schematically in (31).

(31) $\ldots V_{M}\ [_{CP}\ [_{ForceP}\ \text{Force}_{[M]}\ [_{FinP} + \text{Fin}\ [_{IP}\ (DP)\ [_{MoodP}\ [V+T+M]\ [_{TP} \ldots]\ ]\ ]\ ]\ ]\ ]$

identification

The main difference between obligatory and polarity subjunctives therefore is the interpretability of the modal feature on the Force head: it is uninterpretable with obligatory subjunctives, but interpretable with polarity subjunctives. In the case of the polarity subjunctive, the interpretability of the modal feature on Force fundamentally relies on the syntax-discourse interface since the speaker's truth-value inference of the complement clause depends on the wider discourse. Unlike French, there is no register that allows subjunctive clauses (nor *for-to* infinitives) as complements to negated epistemics in English, as illustrated in (31).

(32) * I don't believe that John be the best candidate for the job.

5.3 Acquisition Task Predictions

The learnability task for the polarity subjunctive ultimately depends on whether the syntactic representations for the obligatory subjunctive are already in place, since, without these representations, the acquisition of the polarity subjunctive cannot proceed. Nevertheless, assuming that the L2 learner has successfully reconfigured their interlanguage grammar to acquire obligatory subjunctives, they must revalue the modal feature in ForceP from (uninterpretable) [uM] to (interpretable) [M]. Here, the interpretable feature, [M], on Force is contingent on the syntax–discourse interface, since a speaker's commitment to the truth-value of the CP depends on the wider discourse.

We predict that learners will only be able to revalue the modal feature in ForceP to an interpretable modal feature, [M], at intermediate proficiency levels, once the syntactic representations for obligatory subjunctives are in place. We assume that if learners do not demonstrate accurate use of obligatory subjunctives, any mood alternation in polarity contexts is not a systematic indication of successful acquisition. Even following successful acquisition of obligatory subjunctives, its felicitous use and/or production in polarity contexts is likely to still be variable in light of previous research concerning the difficulty of constructions implicating discourse-contextual information external to the grammar even at an advanced developmental stage (Sorace, 2011).

The success of feature reassembly, however, ultimately relies on abundant and unambiguous evidence in the input. Slabakova (2015, p. 680) argues that the input plays a crucial role in "setting parameter values within the generative framework" and that successful acquisition depends, in large part, on the quality and quantity of exposure. Constructions that are not salient and/or frequent in the input are therefore expected to be more problematic for L2 learners. This input is likely to be a complicating factor in the L2 acquisition of the French subjunctive given its lack of salience and infrequency (O'Connor DiVito, 1997; Poplack, Lealess and Dion, 2013).

5.4 Studies on L2 Acquisition of Polarity Subjunctive

As we discussed in Chapter 4, research has shown that the French subjunctive can be a source of difficulty for many L2 learners, but that such difficulties can eventually be overcome, at least in the case of obligatory subjunctives (Howard, 2008; Bartning, Lundell and Hancock, 2012; Ayoun, 2013; McManus and Mitchell, 2015). The factors contributing to these difficulties include, but are not limited to, crosslinguistic influence, proficiency and exposure to the target language, among others. Limited research, however, has directly investigated the use of the French subjunctive in variable contexts; the only exception is a study by McManus and colleagues (Mcmanus *et al.*, 2014; McManus and Mitchell, 2015). Note that by variable contexts, we refer here to polarity subjunctives, as defined in the section above. The current section aims to present an overview of these studies, focusing not only on L2 French but also on L2 Spanish, given the paucity of research across both languages.

One of the earliest studies to examine the L2 acquisition of the polarity subjunctive was Borgonovo and Prévost (2003). They investigated knowledge of the polarity subjunctive in complement clauses in Spanish by French-speaking learners. Although both French and Spanish express the polarity subjunctive, the contexts vary. In Spanish, epistemic verbs (e.g., *creer* 'believe'), perception verbs (e.g., *ver* 'see') and communication verbs (*decir* 'say') trigger the polarity subjunctive. In French, however, the polarity subjunctive is limited to the complement clauses of negated epistemic verbs; only the indicative is allowed with perception and communication verbs. Findings from a truth-value judgment task revealed that the most advanced learners performed within the range of the native speakers, allowing the subjunctive but rejecting the indicative with predicates that do not allow the subjunctive in their L1 French. In other words, the L2 learners demonstrated target-like behaviour, despite L1/L2 differences in licensing conditions.

Iverson, Kempchinsky and Rothman (2008) further explored the L2 acquisition of the polarity subjunctive in Spanish, but this time with English-speaking learners. However, unlike the Borgonovo and Prévost study where the L1 French grammar contained the grammatical structure, albeit with a more limited subset of verbs, the L1 English grammar in this study only allowed indicative complements with these verb types. As Iverson and colleagues argue, this required the L2 grammar to revalue the modal feature on the Force head from uninterpretable to interpretable. As we have previously stated, the modal feature must be valued as interpretable on the Force head since this feature links to the wider discourse context and specifies whether the speaker or the subject of the matrix clause is committed to the truth-value of the embedded proposition. Results from a scalar grammaticality judgment task showed that the advanced L2 learners' knowledge of the polarity subjunctive "approach[ed] (and sometimes equal[ed]) those of the [native speaker] group", suggesting target-like behaviour (Iverson, Kempchinsky and Rothman, 2008, p. 156). Although the intermediate L2 group did not exhibit native-like behaviour, they did demonstrate a tendency towards it. The authors therefore concluded that interface conditions represent a source of L2 delays and variability, but that target-like acquisition of syntax-discourse interface conditions is ultimately attainable.

Task modality has also been found to modulate the type of knowledge that L2 learners (and heritage speakers) exhibit during language comprehension. For example, van Osch, Hulk, Aalberse and Sleeman (2018) compared proficiency-matched heritage and L2 speakers of Spanish with Dutch as their dominant language on their implicit and explicit knowledge of the subjunctive. The results revealed that the heritage speakers outperformed the L2 speakers on the implicit (elicited oral production) task with respect to their knowledge of the obligatory subjunctive (morphosyntactic) condition and the negated epistemics (syntax-discourse) condition, whereas the reverse was true for the explicit (acceptability judgment) task. The authors therefore argued that such findings underscored the importance of using different task types in bilingual research.

Unlike in Spanish, the polarity subjunctive in French has attracted relatively little attention in the L2 acquisition literature. To our knowledge, one of the few studies to investigate subjunctive use following negated epistemics is McManus and Mitchell (2015). As discussed in Section 4.4, McManus and Mitchell analysed data from two (oral and written) production tasks and a grammaticality judgment task in order to examine how subjunctive use developed over the course of a residence abroad experience. Although residence abroad only had a limited effect on L2 subjunctive use, the authors noted that the learners performed more accurately with affirmative triggers than with conjunctions and negated epistemics. These findings echo those of Iverson *et al.*'s study, showing more accurate subjunctive use in non-variable contexts (i.e., with affirmative triggers) than variable contexts (i.e., with negated epistemics). Note, however, that unlike the Iverson and colleagues' study, McManus and Mitchell did not manipulate commitment to the truth-value of the embedded proposition. Instead, the authors simply assumed that the subjunctive was the grammatical form and the indicative the ungrammatical form in such contexts.

It is apparent from the discussion presented above that the polarity subjunctive represents a source of persistent difficulty and, in some cases, developmental delay for many L2 learners, but that this initial delay can eventually be overcome at more advanced stages of proficiency. Furthermore, research has shown that task type can significantly modulate the type of knowledge that speakers activate during language comprehension and production, such that L2 learners are more likely to demonstrate target-like knowledge in tasks that focus on form rather than function, such as acceptability judgment tasks. However, what is most clear from the above discussion is the paucity of research on the polarity subjunctive in French. The current study thus seeks to address this by investigating English-speaking L2 learners' knowledge of the discourse-pragmatic constraints governing the French polarity subjunctive.

5.5 Processing of Counterfactual Thought and Negation

As discussed in the previous section, research has shown that the polarity subjunctive can be a source of difficulty for many L2 learners, particularly at earlier stages of development (Iverson, Kempchinsky and Rothman, 2008; McManus and Mitchell, 2015). Although no study to date has investigated online sensitivity to the polarity subjunctive, a number of studies have examined the processing of counterfactual thought among the L1 population (cf., Kulakova and Nieuwland, 2016b for a review of these studies).

Counterfactual thought enables individuals to entertain possible worlds they know to be false. This is similar in many respects to the processing of the polarity subjunctive. As we have stated previously, mood use under negated epistemics depends on a speaker's commitment to the truth-value of the embedded clause. Although the subjunctive is typically the expected form in the embedded complement clauses of negated epistemics, the indicative can also be used depending on the speaker's (or the subject of the matrix clause's) commitment to the truth-value of the embedded clause. If the speaker (or the subject of the matrix clause) is committed to the truth-value of the embedded clause, then the indicative (and also the subjunctive) is licensed, whereas if the speaker is not committed to the truth-value of the embedded clause, then the subjunctive is licensed, but not the indicative.

Research on the processing of counterfactual thought therefore has the potential to further expand our understanding of how the subjunctive is processed in polarity contexts among both the L1 and L2 population. With this in mind, the current sub-section aims to provide a brief overview of previous studies looking at the processing of counterfactual thought.

Cognitive accounts of counterfactual thought (Fauconnier, 1994; Johnson-Laird and Byrne, 2002) argue that counterfactuals are cognitively demanding because they elicit two incompatible representations (Kulakova and Nieuwland, 2016b). For example, the proposition, "If I had wings, then I would be able to fly", expresses the following two representations: (1) that the suppositional but factually false state of the speaker having wings and being able to fly, while also expressing (2) that the speaker does not have wings and therefore relies on conventional modes of transportation (Kulakova and Nieuwland, 2016b)

Previous research has shown that when readers come across an anomalous word that does not fit into the sentence, they are more likely to fixate on this word for longer and make more backward regressions compared with a word that is compatible with the sentence context, which suggests that speakers are sensitive to discourse-level incongruencies (e.g., Ni *et al.*, 1998; Braze *et al.*, 2002; Rayner *et al.*, 2004). With this in mind, Ferguson and Sanford (2008) examined counterfactual processing using an eye-tracking during reading task within an anomaly detection paradigm. In this study, participants were presented with two sentences establishing factual and counterfactual scenarios. The first sentence determined whether the context was factual (e.g., *If cats were hungry* …) or counterfactual (e.g., *If cats were vegetarians*…). The second sentence described a hypothetical situation that was "either consistent or inconsistent with the context, dependent on a critical word (e.g., feeding a cat a bowl of fish vs. carrots)" (Black, Williams and Ferguson, 2018, p. 1444). This resulted in a 2 × 2 design with four conditions: real world inconsistent, real world consistent, counterfactual world inconsistent or counterfactual world consistent. Results revealed that readers were able to quickly integrate the novel counterfactual world, with longer reading times at the critical word that did not match the described counterfactual world than when the critical word matched the counterfactual world. However, detection of counterfactual world inconsistencies was delayed in comparison to factual world inconsistencies, since readers were required to restrict real-world knowledge before detecting counterfactual inconsistencies.

Although counterfactuals enable individuals to discuss topics that extend beyond purely veridical propositions, it is possible that this flexibility comes at a cost during processing (Kulakova and Nieuwland, 2016b). Previous research focusing on online sentence comprehension has shown that counterfactual antecedents involving new factual meaning rather than existing factual meaning result in longer reading times (Rayner, 1998, 2009). This research has found that individuals immediately verify counterfactual events with the previous discourse for consistency (Stewart, Haigh and Kidd, 2009; Haigh and Stewart, 2011; Ferguson, 2012). Furthermore, counterfactual contexts appear to be quickly integrated into the discourse representation, to the extent that they influence the comprehension of the subsequent narrative (Ferguson, 2012; Ferguson and Cane, 2015). This process is argued to rely heavily on working memory resources (Ferguson and Cane, 2015).

Recent research has found that pragmatic skill also influences the online comprehension of counterfactual antecedents. A study by Kulakova and Nieuwland (2016a), using electroencephalography (EEG) data, investigated whether speakers were able to integrate overt cues (subjunctive mood, e.g., *If I loved you then*) quickly enough to understand the expected counterfactual meaning and whether this ability was modulated by pragmatic skill (that is, the ability to integrate knowledge of the social-communicative function of language in everyday life). The study revealed two key findings: (1) that counterfactuality is rapidly integrated during language comprehensions and moderates online expectations in line with factual knowledge and (2) that this ability to reduce online expectations was influenced by an individual's ability to understand the communicative intentions of others. This is, "individuals who are better at understanding the communicative intentions of other people are more likely to reduce knowledge-based expectations in counterfactuals" (Kulakova and Nieuwland, 2016a, p. 822).

5.6 Research Questions and Predictions

In response to limited existing research, the current study aims to investigate the influence of discourse-pragmatic constraints on the acquisition and processing of the polarity subjunctive in

L2 French at different levels of proficiency and residence in a French-speaking environment. With this in mind, we advance the following research questions and related predictions:

- Do English-speaking L2 learners of French exhibit offline and online sensitivity to the syntactic and discourse/pragmatic constraints of the subjunctive in polarity contexts?

- To what extent is this sensitivity influenced by L1/L2 differences in
    - proficiency and
    - residence in a French-speaking country?

Based on the literature reviewed in the previous sections, sensitivity is operationalised as follows. In particular, we expect reading measures to be longer (or higher in the case of probabilities) for the ungrammatical condition (Indicative-False) than the other three grammatical conditions (Indicative-True, Subjunctive-True and Subjunctive-False). Reading times may also be longer for the Indicative-True condition, regardless of the discourse-pragmatic constraints governing the polarity subjunctive, in light of the observation that the subjunctive is often the preferred form in written language since it is seen as a marker of formality (Müller and Elsass, 1985; O'Connor DiVito, 1997; Lenoble-Pinson and Grevisse, 2009; Jeppesen Kragh, 2010). Given that the polarity subjunctive requires the integration of discourse-pragmatic information, it is likely that sensitivity will be found predominately in the later reading measures, namely go-past time, total reading time and regression-in and regression-out probabilities. These measures in particular have been shown to reflect integration difficulties and higher-level (e.g., discourse) processing (Rayner and Pollatsek, 1989; Liversedge, Paterson and Pickering, 1998; Rayner *et al.*, 2004).

As discussed in Section 5.2.2, the learnability task for the polarity subjunctive ultimately depends on whether the syntactic representations for the obligatory subjunctive are already in place. In Study 1, we established that while L2 learners' online knowledge of the subjunctive was not always target-like, their offline knowledge was. We interpreted such a finding as indicating that the L2 speakers had acquired the relevant syntactic representations, but that cognitive resource constraints (e.g., working memory restrictions) impeded target-like processing patterns.

With this in mind, we anticipate that L2 learners will demonstrate an asymmetry between their offline and online knowledge of the polarity subjunctive, with more target-like sensitivity in the offline than the online data, as was the case in Study 1. Furthermore, we predict that at lower proficiency levels, L2 learners will exhibit limited to no sensitivity to the ungrammaticality of the indicative, let alone to the discourse-pragmatic constraints, in their online processing and will display longer reading times for the subjunctive than indicative sentences, regardless of their commitment to the truth-value of the embedded proposition. In other words, lower proficiency L2 learners will initially treat subjunctive complements to negated epistemics as ungrammatical given that the L1 (British English) only permits indicative complements in this context.

However, as proficiency and exposure to the target language increase, we anticipate that L2 learners will treat negated epistemic predicates as licensing only subjunctive complements, regardless of the discourse-pragmatic constraints, in a similar manner to obligatory contexts, with longer reading measures for indicative than subjunctive sentences to reflect the ungrammaticality of the indicative in the interlanguage grammar. It is likely that such a stage will be reflected initially in the learners' offline knowledge and only at a later stage of proficiency and exposure, in their online knowledge.

Finally, it is at the most advanced stages of proficiency and highest levels of exposure to the target language that we anticipate that L2 learners will exhibit target-like knowledge of the discourse-pragmatic constraints governing the polarity subjunctive, with longer (or higher in the case of probabilities) reading measures for the ungrammatical condition (Indicative-False) than the other three grammatical conditions (Indicative-True, Subjunctive-True and Subjunctive-False). Again, it is plausible that offline knowledge will precede online knowledge.

(33) Summary of predictions:

- Hypothesis 1: L2 learners should be more sensitive to the discourse-pragmatic constraints in the offline data than the online data, but this sensitivity is not guaranteed.
- Hypothesis 2: As proficiency increases, L2 learners should exhibit increased sensitivity to the discourse-pragmatic constraints.
- Hypothesis 3: Exposure to (naturalistic) input, through residence abroad, should improve sensitivity to discourse-pragmatic constraints.

5.7 Method

5.7.1 Participants

Participants were identical to those included in Study 1. To recap, we tested a total of 75 participants, including 30 L1 speakers of French and a test group of 45 (English-speaking) L2 learners of French.

5.7.2 Eye-Tracking

5.7.2.1 Materials

In Study 2, we recorded participants' eye movements while they read sentences, as in Table 5.1. Prior to reading each sentence, participants were presented with an image designed to establish the speaker's commitment to the truth-value of the embedded proposition. All test items were bi-clausal sentences and followed the same underlying syntactic structure up until the critical region (i.e., the verb in the embedded clause), consisting of a singular third-person [+ animate] noun, a negated epistemic verb (i.e., *penser* 'to think' or *croire* 'to believe') and the complementiser *que* 'that' in the matrix clause followed by a singular third-person [+ animate] noun, the critical verb (i.e., *être* 'to be', *faire* 'to do', or *avoir* 'to have') marked with the relevant mood morphology (indicative or subjunctive) and the subsequent continuation in the embedded clause, including a sentence-final disambiguating region (i.e., *mais moi je crois que si* 'but I think so' / *et je ne le pense pas non plus* 'and I don't think so either').

Table 5.1 Example of test sentences in polarity condition.

Mood	Truth	Example
Indicative	False	* Juliette ne pense pas que le chat **a** les yeux verts et je ne le pense pas non plus.
		'Juliette does not think that the cat has green eyes and I don't think so either.'
Indicative	True	Juliette ne pense pas que le chat **a** les yeux verts mais moi je pense que si.
		'Juliette does not think that the cat has green eyes but I think it does.'
Subjunctive	False	Juliette ne pense pas que le chat **ait** les yeux verts et je ne le pense pas non plus.
		'Juliette does not think that the cat has green eyes and I don't think so either.'
Subjunctive	True	Juliette ne pense pas que le chat **ait** les yeux verts mais moi je pense que si.
		'Juliette does not think that the cat has green eyes but I think it does.'

For each test item, we manipulated two within-subject factors. The first was commitment to the truth-value of the embedded clause, which had two levels/conditions, and for which we selected thirty-two versions per condition from an initial set of forty-one. Within each level/condition, we manipulated the speaker/reader's commitment to the truth-value of the embedded proposition with a picture and further reinforced it with the aforementioned disambiguating region.

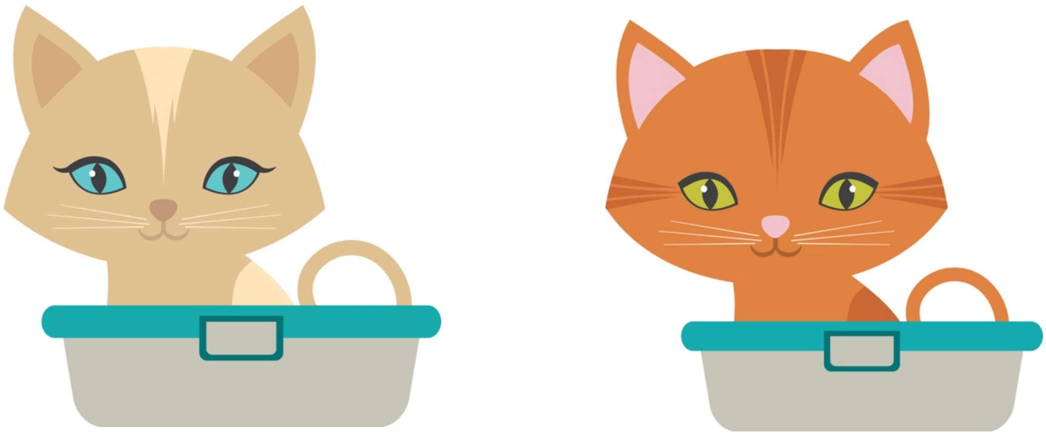


Figure 5.1 Example of pictures

In the +commitment condition, the picture had the property or conveyed the action expressed by the predicate in the embedded clause, thus contradicting the opinion of the subject of the matrix clause (see the rightmost image in Figure 5.1). The +commitment condition also included the disambiguating region/clause, *mais moi je crois que si,* ‘but I think so’, reinforcing the speaker/reader’s commitment to the truth-value of the embedded clause.

In the –commitment condition, the picture did not have the property or convey the action expressed by the predicate in the embedded clause, thus supporting the opinion of the subject of the matrix clause (see the leftmost image in Figure 5.1). The –commitment condition also included the disambiguating region/clause, *et je ne le crois pas non plus* ‘and I don’t think so either’, highlighting the speaker/reader’s lack of commitment to the truth-value of the embedded clause.

The second factor manipulated was verbal mood (i.e., indicative and subjunctive) in the embedded clause, for which we created thirty-two versions for each condition. Exemplar test items are provided in Table 5.1.

Table 5.2 Grammaticality of each mood according to truth-value

	True [+commitment]	False [–commitment]
Indicative	grammatical	ungrammatical
Subjunctive	grammatical	grammatical

Table 5.2 outlines the (un)grammaticality of each mood based on commitment to the truth-value of the embedded proposition. Henceforth, and for ease of exposition, the

abbreviation, *True,* will refer to conditions where the speaker/reader is pushed towards agreeing with the truth-value of the embedded proposition, whereas *False* will indicate conditions where the speaker/reader is encouraged to reject the truth-value of the embedded proposition.

5.7.2.2 Lexical Properties of Test Items

Table 7.3 in Appendix A.3.1 summarises the lexical characteristics of the critical regions in this experiment. Recall that in Section 2.2.1, we observed that three lexical properties (i.e., frequency, length and predictability) in particular strongly influenced reading times. For this reason, eye-tracking researchers typically attempt to match the critical regions across conditions and items on frequency, length and predictability. The ANOVA results in Table 7.4 in Appendix A.3.1 revealed significant differences between conditions for critical region length and zipf frequency (film subtitles and books). We argue that these differences were, to a certain extent, unavoidable due to the fact that the French subjunctive is an infrequent linguistic structure in the input where the subjunctive forms are typically longer in characters than indicative forms. For this reason, and to address any potential confounds, we included critical region length and/or frequency as control variables in our statistical models.

5.7.2.3 Piloting

To ensure that the test items exhibited the expected range of acceptability and rule out any unexpected differences in the material, we ran three pre-tests, each completed by eight self-identified native speakers of French who did not participate in the main study.

The first pre-test was a norming study on Amazon Mechanical Turk, designed to ensure that the images expressed the intended action or state. Eight self-identified native speakers of French took part in a picture-sentence verification task for a small compensation. They were asked to assess whether the sentence accurately matched the action or state shown in the picture, on a scale from 1 (completely unrelated) to 7 (perfectly matched). A set of 41 items were initially created using the embedded clause of the main experiment's test items. For each item, two versions were created; version A was an accurate (TRUE) description of the picture and version B an inaccurate (FALSE) description. The items were split across two lists in a Latin-square design. True picture/sentence combinations attracted a mean rating of 6.24 (SD = 0.94) and false picture/sentence combinations a mean rating of 1.65 (SD = 1.36). The estimates from a mixed-effects linear regression model, presented in Table 5.3 below, revealed a significant difference between true and false sentences.

Table 5.3 Mixed-effects models for first pre-test (polarity study)

	b	*SE*	*t*
(Intercept)	3.95	0.12	34.25
Truth	2.30	0.07	33.85

In order to rule out any unintended grammaticality differences in the material within the embedded clause, a second pre-test was conducted on Amazon Mechanical Turk. Eight self-identified native speakers of French took part in an acceptability judgment task for a small compensation. Their task was to rate the grammaticality of simple affirmative clauses/sentences, on a scale from 1 (completely unacceptable) to 7 (completely acceptable). The sentences received a mean rating of 6.74 ($SD = 0.78$).

Table 5.4 Mean ratings from picture-sentence verification task

	Indicative	Subjunctive
True	5.60 (1.70)	5.72 (1.65)
False	5.99 (1.51)	6.00 (1.36)

A third and final pre-test was conducted via Amazon Mechanical Turk in order to assess whether the test items exhibited the desired range of acceptability. This task was designed to emulate the main experiment; the only difference was that participants' eye movement were not tracked and no time constraints were imposed. Fourteen native speakers of French took part in a picture-sentence verification task for a small compensation. While their task was to evaluate whether the sentence matched the action or state shown in the picture, on a scale from 1 (completely unacceptable) to 7 (completely acceptable), participants were told to base their final judgment on the grammaticality of the sentence. A mixed-effects linear regression model (see Table 5.4) revealed a significant difference between true and false sentences, but not between indicative and subjunctive sentences.

Table 5.5 Mixed-effects model for third pre-test (polarity study)

	b	*SE*	*t*
(Intercept)	5.80	0.26	22.10
Mood: Subjunctive vs. Indicative	0.01	0.06	0.24
Truth: True vs. False	-0.17	0.06	-2.88
Mood × Truth	0.03	0.06	0.51

Note: Shaded cells denote statistical significance ($p < .05$).

5.7.2.4 Design

Of the initial 41, a set of 32 experimental items were selected.[31] The items were pseudo-randomised and distributed across four presentation lists in a Latin-square design, interspersed with an equal number of filler items. Half of all trials were followed by a yes-no comprehension question. The comprehension question probed the participant's commitment to the truth-value of the embedded clause. The main experiment was divided into two blocks to allow participants to take a break if needed. Six practice items were included before the main experiment to familiarise participants with the procedure.

5.7.2.5 Apparatus

Participants' eye movements were recorded using the SR Research EyeLink 1000 Plus system with a sampling rate of 1000 Hz as they read the stimuli on a 24-inch CRT monitor with a refresh rate of 60 Hz. The sentences were displayed in black monospaced font of approximately 22 pt on a white background. Although viewing was binocular, eye movements were recorded from the right eye. Participants were seated approximately 90 cm from the monitor and head movements were minimized using a chin and forehead rest.

5.7.2.6 Procedure

Both studies were identical with respect to stimuli. The only difference between the two studies was the type of secondary task used. The order in which the two parts of the experiments took place was fixed and thus not counterbalanced, with the judgment task always following the comprehension task.[32]

In Study 2a, participants were instructed to look at the contextualising picture (for a maximum of 15 seconds), silently read each sentence for comprehension at their normal reading

---

[31] The 32 experimental items were chosen for inclusion on the basis on the grammaticality of the embedded clause, as tested in the second pre-test. In order words, we selected the sentences with the most highly rated embedded clauses.

[32] We acknowledge that this will have led to repetition effects and quite possibly shorted overall reading times in the judgment task due to repeated exposure to experimental stimuli. However, we consciously adopted this order since the comprehension task was designed to reflect more closely natural reading, allowing participants to read for comprehension rather than for form. In this sense, it was designed to be the more implicit of the two tasks. Had the judgment task preceded the comprehension task, it is likely that the participants would have focused more on grammatical form in the comprehension task, which would have contradicted the overall goal of the comprehension task.

rate (for no more than 10 seconds) and then answer a question (within 15 seconds) on 50% of all trials. Comprehension questions focused on whether the contextualising picture matched the sentence and required a yes-no push-button response.

In Study 2b, participants were asked to look at the contextualising picture (for a maximum of 15 seconds), silently read each sentence for a maximum of 20 seconds and then judge how natural the sentence sounded by selecting a number between 1 and 7, where '1' indicated that the sentence was completely unacceptable and '7' completely acceptable. An acceptability judgment task was included in this study to probe speakers' offline knowledge of the subjunctive.

Both experiments started with a nine-point calibration procedure followed by six practice trials and a total of 32 experimental trials, which were presented in an individually randomized order and intermixed with 32 filler items. At the start of each trial, a fixation point appeared at the location of the first letter of the sentence. Once the participant's fixation on this point was stable, the picture was presented. Where necessary, recalibration was performed to compensate for any drift in calibration. We aimed for a mean calibration of 0.5 degrees of visual angle. The entire experiment lasted approximately 10 minutes in total.

5.7.2.7 Regions of Interest

Sentences were divided up into three regions of interest: the predicate in the matrix clause, the embedded verb and the subsequent condition. Eye movements were analysed at the critical region (the embedded verb in all conditions) and the spillover region (one or two words of up to a total of 20 characters).[33]

---

[33] In the experimental stimuli, the word following the critical region varied considerably in length. In order to minimise this variation and ensure a certain level of consistent across test items, we extended the spillover region to include up to three words. We recognise that the variation in the spillover region length is a confounding factor in this study and controlled for it post-hoc by including region length (in characters) as a control variable in the mixed-effects models.

Table 5.6 Regions of interest in the polarity condition.

		matrix predicate		pre-target	target	post-target	subsequent continuation
Indicative True	Quentin	ne pense pas	que	Julien	est	en colère	mais moi je pense que si
Subjunctive True	Quentin	ne pense pas	que	Julien	soit	en colère	mais moi je pense que si
Indicative False	Quentin	ne pense pas	que	Julien	est	en colère	et je ne le pense pas non plus
Subjunctive True	Quentin	ne pense pas	que	Julien	soit	en colère	et je ne le pense pas non plus

5.8 Results: Study 2a (Comprehension)

5.8.1 Data Preparation and Analyses

Following standard procedure in psycholinguistic studies (Rayner and Pollatsek, 1989), fixations shorter than 80 ms were merged with the neighbouring fixation if the two fixations were within one character of each other. All other fixations shorter than 80 ms or exceeding 800 ms were excluded from the dataset (Rayner 1998; Rayner et al. 2004). Fixations occurring immediately before or after a blink were also removed from the analysis. Finally, we omitted fixations that were more than three standard deviations from the mean of each dependent variable.

To correct the skewed distribution of the raw data and approximate normal distribution, fixation data was log transformed. For the continuous variables, mixed-effects linear regression models were conducted using the *lmer* function of the *lme4* package (Bates *et al.*, 2015) in the R environment (R Development Core Team, 2014). For binary variables[34] (regression probabilities), mixed-effects logistic regression models were computed using the *glmer* function of the aforementioned package.

Where possible, we fitted each model using the 'maximal' random effects structure (i.e., slopes for each fixed effect across item and subject) that converged (Barr *et al.*, 2013). In many cases, the maximal random effects structure had to be cut down due to non-convergence or a singular fit (i.e., perfect or near perfect correlations in the random structure). If a model did not converge or had a singular fit, parameters were systematically excluded from the random effects

[34] In this thesis, we only report the descriptive statistics for skipping probability. See footnote 20 on page 59 for the justification behind this decision.

structure based on the level of variance, such that parameters with the lowest variance were excluded first until the model converged. Absolute *t* values (or *z* values in the case of binary variables) exceeding ± 1.96 were analysed as significant.

In light of our earlier discussion concerning the comparative fallacy, L1 and L2 fixation data were analysed separately. For the L1 data, mixed-effect models were computed for each fixation measure to examine whether there was a significant difference between mood (indicative and subjunctive), commitment to truth-value ([-commitment] and [+commitment]) and/or their interaction. Both mood and truth-value were inputted as categorical fixed effects and sum contrasts were coded. For mood, we explored the difference between indicative and subjunctive sentences (indicative -1, subjunctive 1). For truth-value, we examined the difference between [+commitment] and [-commitment] ([-commitment] -1, [+commitment] 1). The mixed-effect models for the L2 data were almost identical to the L1 models. The only exception was the inclusion of two fixed effects (i.e., proficiency (centred) and residence in a French-speaking country) in the L2 models.

To control for frequency and/or length confounds, we inputted at least one control variable as a continuous fixed effect. For global measures, we chose sentence length (centred) in characters and for local measures, region frequency or length (centred) in characters. We originally planned to include both region frequency and length for the critical region. However, this would have led to multicollinearity of variables between frequency and length. Consequently, we only selected region frequency in the critical region analyses and region length in the spillover region analyses. We also entered trial index as a continuous fixed factor to avoid any confounds relating to the position of each test item in the experiment. Although the addition of these parameters as control variables were important, we will not discuss them in any more depth.

Table 5.7 (repeated for ease of reference) illustrates that both the indicative and subjunctive forms were considered grammatical under the [+commitment] condition, but that only the subjunctive under the [–commitment] condition in the truth-value factor. As such, we expected sensitivity to the discourse constraints of the polarity subjunctive to manifest in the interaction between mood and truth-value. However, considering that only one out of the four conditions was classed as ungrammatical, it is indeed possible that such an effect may not show up in the inferential analyses. For this reason, we conducted a series of pairwise post-hoc comparisons using the *emmeans* package in the R environment. This data will be presented alongside the discussion of the mixed-effects models. Note that we will only present significant differences (if found) between the ungrammatical condition (Indicative-False) and any of the three grammatical conditions.

Table 5.7 Grammaticality of each mood according to truth-value

	True [+commitment]	False [–commitment]
Indicative	grammatical	ungrammatical
Subjunctive	grammatical	grammatical

Table 5.8 Mixed-effects models for L1 global in comprehension

		b	*SE*	*t*
Mean Fixation Count	(Intercept)	13.14	0.53	24.63
	Mood: Subjunctive vs. Indicative	-0.21	0.15	-1.39
	Truth: True vs. False	-0.17	0.17	-1.02
	Sentence Length (Centred)	0.63	0.18	3.45
	Trial Number	-0.05	0.01	-6.16
	Mood × Truth	0.07	0.15	0.47
Mean Fixation Duration	(Intercept)	5.38	0.02	218.26
	Mood: Subjunctive vs. Indicative	0.01	0.01	2.02
	Truth: True vs. False	-0.01	0.01	-2.61
	Sentence Length (Centred)	-0.02	0.01	-2.94
	Trial Number	0.00	0.00	-3.40
	Mood × Truth	0.00	0.00	-0.55
Total Sentence Reading Time	(Intercept)	7.40	0.14	52.93
	Mood: Subjunctive vs. Indicative	-0.01	0.02	-0.61
	Truth: True vs. False	-0.05	0.02	-2.45
	Sentence Length (Centred)	0.05	0.02	2.65
	Trial Number	-0.01	0.00	-8.50
	Mood × Truth	0.00	0.02	0.05

Note: Shaded cells denote statistical significance ($p < .05$).

Table 5.9 Mixed-effects models for L2 global in comprehension

	Mean Fixation Count			Mean Fixation Duration			Total Sentence Reading Time		
	b	*SE*	*t*	*b*	*SE*	*t*	*b*	*SE*	*t*
(Intercept)	17.86	1.99	8.96	5.56	0.06	91.31	7.98	0.42	18.81
Mood: Subjunctive vs. Indicative	-0.48	0.60	-0.81	0.01	0.01	0.40	-0.04	0.04	-1.08
Truth: True vs. False	-0.36	0.61	-0.60	0.01	0.01	0.79	-0.03	0.04	-0.71
Proficiency (Centred)	1.54	2.11	0.73	0.14	0.06	2.16	-0.11	0.46	-0.24
Residence (in French-Speaking Country)	1.32	2.26	0.59	-0.15	0.07	-2.15	-0.02	0.48	-0.04
Sentence Length (Centred)	0.40	0.20	1.98	0.00	0.01	-0.30	0.03	0.01	1.86
Trial Number	-0.08	0.01	-7.83	0.00	0.00	-2.55	-0.01	0.00	-9.50
Mood × Truth	-0.14	0.60	-0.23	-0.02	0.01	-1.65	-0.05	0.04	-1.19
Mood × Prof.	-0.50	0.63	-0.80	0.01	0.02	0.35	-0.05	0.04	-1.12
Truth × Prof.	0.21	0.63	0.34	0.03	0.02	1.87	0.00	0.04	0.05
Mood × Residence	-0.11	0.70	-0.16	-0.01	0.02	-0.31	0.01	0.05	0.12
Truth × Residence	-0.45	0.71	-0.64	-0.02	0.02	-0.97	-0.01	0.05	-0.23
Proficiency × Residence	-1.06	2.47	-0.43	-0.17	0.08	-2.19	-0.09	0.52	-0.18
Mood × Truth × Prof.	-0.42	0.63	-0.68	-0.03	0.02	-2.06	-0.04	0.04	-1.06
Mood × Truth × Residence	0.57	0.70	0.81	0.01	0.02	0.49	0.06	0.05	1.29
Mood × Prof. × Residence	0.08	0.75	0.11	-0.01	0.02	-0.30	0.02	0.05	0.31
Truth × Prof. × Residence	-0.58	0.75	-0.77	-0.03	0.02	-1.49	-0.01	0.05	-0.27
Mood × Truth × Prof. × Residence	0.88	0.75	1.17	0.03	0.02	1.53	0.08	0.05	1.71

Note: Shaded cells denote statistical significance ($p < .05$)

5.8.2 Global Measures

5.8.2.1 L1 Group

To examine the overall effect of mood and its interaction with truth-value on reading behaviour, we calculated the following global measures: mean fixation duration, mean fixation count and total sentence reading time. For each measure, observed means are presented in Table 5.10 and model parameters in Table 5.8. Estimated means were used to further interpret the output of the mixed-effects models.

Table 5.10 Raw means (standard deviations) for L1 global in comprehension

		Total Sentence Reading Time	Mean Fixation Count	Mean Fixation Duration
Indicative	False	2224.6 (1563.68)	11.73 (4.13)	212.56 (35.21)
	True	1943.87 (1430.24)	10.76 (3.93)	209.78 (34.52)
Subjunctive	False	2123.84 (1491.77)	11.5 (4.17)	218.11 (39.1)
	True	1866.35 (1329.5)	10.56 (4.01)	212.94 (34.97)

There was a significant main effect of mood for mean fixation duration, but not for mean fixation count or total sentence reading time. In particular, mean fixation duration was significantly longer for subjunctive than indicative sentences. There was, however, a significant main effect of truth-value for mean fixation duration and total sentence reading time, but not for mean fixation count. In particular, false sentences attracted longer mean fixation durations and total sentence reading times than true sentences. None of the global measures exhibited a significant interaction between mood and truth-value.

5.8.2.2 L2 Group

To examine L2 learners' sensitivity to mood and the influence of truth-value, proficiency and residence abroad, we calculated the following global measures: mean fixation duration, mean fixation count and total sentence reading time. For each measure, observed means are presented in Table 5.11 and model parameters in Table 5.9.

Table 5.11 Raw means (standard deviations) for L2 global in comprehension

		Total Sentence Reading Time	Mean Fixation Count	Mean Fixation Duration
Indicative	False	4127.61 (2006.43)	15.85 (6.36)	226.49 (30.75)
	True	3620.74 (1822.51)	13.92 (5.05)	225.03 (31.88)
Subjunctive	False	3920.82 (1853.51)	14.96 (5.46)	228.68 (27.48)
	True	3439.51 (1720.64)	13.79 (5.31)	222.72 (31.11)

The model parameters in Table 5.9 did not reveal a significant main effect of mood or truth-value for any of the three global measures in the L2 data. There was, however, a significant three-way interaction between mood, truth-value and proficiency for mean fixation duration. Figure 5.2 visualises this interaction. It shows that that Subjunctive-False sentences consistently attracted longer mean fixation durations than Indicative-False sentences, particularly as proficiency increased. The reverse was true for true sentences.

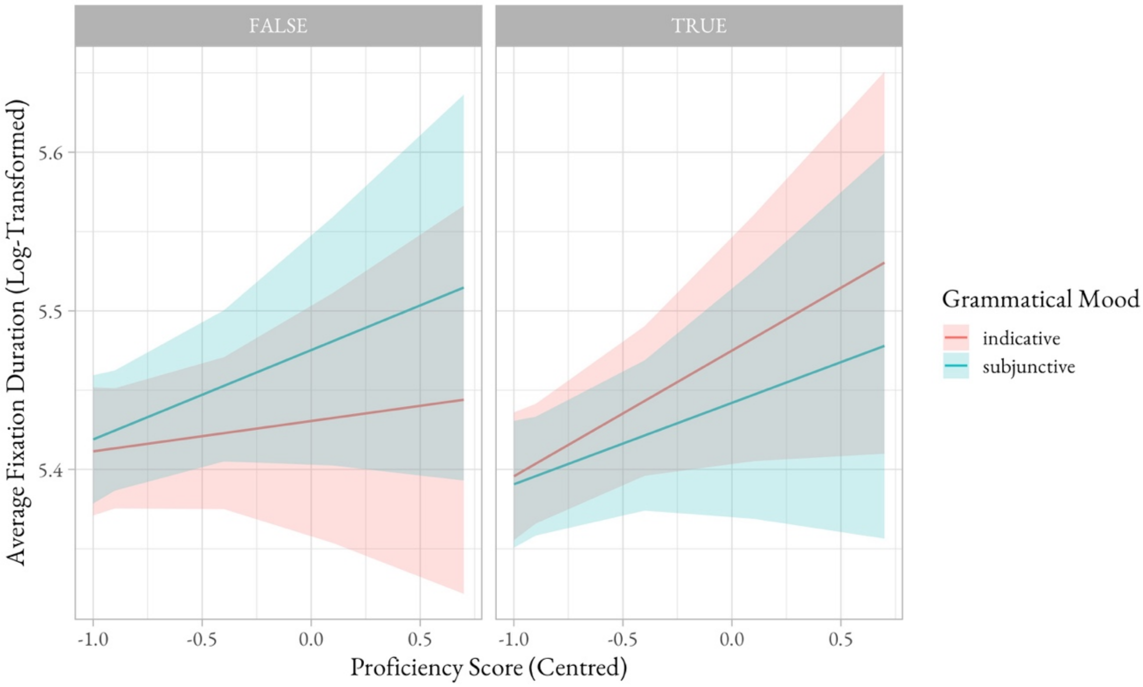


Figure 5.2 Predicted mean fixation duration (comprehension)

Table 5.12 Raw means (standard deviations) for L1 local in comprehension

			First Fixation Duration	Gaze Duration	Go-Past Time	Total Reading Time	Skipping Probability	Regression-In Probability	Regression-Out Probability
Critical Region	Indicative	False	213.51 (110.45)	221.81 (130.99)	233.48 (198.12)	282.14 (296.05)	0.82 (0.38)	0.44 (0.5)	0.02 (0.15)
		True	210.51 (92.8)	210.51 (92.8)	227.95 (120.57)	257.44 (153.64)	0.83 (0.38)	0.36 (0.48)	0.02 (0.15)
	Subjunctive	False	211.49 (79.02)	220.33 (90.14)	228.65 (95.92)	285.09 (191.93)	0.67 (0.47)	0.3 (0.46)	0.05 (0.21)
		True	204.4 (75.45)	205.53 (77.29)	219.34 (91.36)	259.73 (149.4)	0.63 (0.48)	0.2 (0.4)	0.08 (0.27)
Spillover Region	Indicative	False	224.29 (65.61)	318.69 (160.73)	346.14 (168.61)	443.55 (264.26)	0.32 (0.47)	0.14 (0.35)	0.15 (0.36)
		True	209.54 (62.53)	283.72 (127.81)	318.66 (146.15)	410.83 (247.37)	0.21 (0.4)	0.15 (0.36)	0.16 (0.37)
	Subjunctive	False	216.17 (76.1)	291.02 (131.11)	326.56 (132.58)	426.52 (259.13)	0.29 (0.46)	0.14 (0.35)	0.19 (0.39)
		True	213.21 (64.5)	278.59 (133.86)	304.39 (156.3)	387.02 (306.89)	0.26 (0.44)	0.11 (0.31)	0.14 (0.34)

Table 5.13 Raw means (standard deviations) for L2 local in comprehension

			First Fixation Duration	Gaze Duration	Go-Past Time	Total Reading Time	Skipping Probability	Regression-In Probability	Regression-Out Probability
Critical Region	Indicative	False	233.13 (103.9)	238.28 (114.02)	249.53 (115.26)	321.93 (186.98)	0.77 (0.42)	0.35 (0.48)	0.07 (0.26)
		True	234.98 (92.46)	237.09 (96.72)	248.56 (115.85)	305.9 (178.65)	0.69 (0.46)	0.34 (0.47)	0.07 (0.25)
	Subjunctive	False	236.98 (92.26)	245.64 (99.12)	256.55 (112.8)	390.24 (289.29)	0.67 (0.47)	0.43 (0.5)	0.06 (0.25)
		True	238.58 (88.93)	254.86 (113.6)	264.43 (120.97)	379.13 (247.46)	0.61 (0.49)	0.36 (0.48)	0.06 (0.24)
Spillover Region	Indicative	False	237.48 (90.96)	412.15 (291.36)	476.05 (316.88)	726.01 (598.58)	0.25 (0.44)	0.25 (0.44)	0.2 (0.4)
		True	231.48 (94.62)	355.43 (208.48)	422.18 (274.97)	584.35 (414.6)	0.21 (0.41)	0.2 (0.4)	0.21 (0.4)
	Subjunctive	False	246.31 (94.88)	407.89 (270.49)	511.03 (319.9)	739.47 (531.45)	0.25 (0.43)	0.23 (0.42)	0.26 (0.44)
		True	235.25 (86.56)	358.37 (223.5)	417.95 (268.05)	597.56 (479.91)	0.25 (0.43)	0.21 (0.41)	0.22 (0.42)

Table 5.14 Mixed-effects models for L1 local in comprehension

		Critical Region			Spillover Region		
		b	*SE*	*t/z*	*b*	*SE*	*t/z*
First Fixation Duration	(Intercept)	5.23	0.04	116.93	5.37	0.03	191.20
	Mood: Subjunctive vs. Indicative	0.00	0.02	0.11	-0.01	0.01	-0.87
	Truth: True vs. False	-0.01	0.02	-0.90	-0.02	0.01	-1.82
	Sentence Length (Centred)	0.00	0.02	-0.07	-0.01	0.01	-0.59
	Trial Number	0.00	0.00	1.15	0.00	0.00	-1.70
	Mood × Truth	-0.01	0.02	-0.72	0.02	0.01	1.59
Gaze Duration	(Intercept)	5.24	0.05	109.69	5.65	0.04	145.71
	Mood: Subjunctive vs. Indicative	0.01	0.03	0.42	-0.02	0.01	-1.86
	Truth: True vs. False	-0.03	0.02	-1.63	-0.04	0.02	-2.76
	Sentence Length (Centred)	0.00	0.03	0.11	0.17	0.02	8.99
	Trial Number	0.00	0.00	0.95	0.00	0.00	-3.07
	Mood × Truth	-0.01	0.02	-0.84	0.01	0.01	1.06
Go-Past Time	(Intercept)	5.31	0.05	103.93	5.76	0.04	154.78
	Mood: Subjunctive vs. Indicative	0.02	0.03	0.55	-0.02	0.01	-1.53
	Truth: True vs. False	-0.01	0.02	-0.58	-0.05	0.02	-3.19
	Sentence Length (Centred)	0.00	0.03	0.08	0.22	0.02	11.47
	Trial Number	0.00	0.00	0.11	0.00	0.00	-3.73
	Mood × Truth	-0.02	0.02	-1.07	0.00	0.01	-0.35
Total Reading Time	(Intercept)	5.46	0.07	82.55	6.02	0.05	120.98
	Mood: Subjunctive vs. Indicative	0.03	0.03	0.94	-0.03	0.02	-1.74
	Truth: True vs. False	-0.03	0.03	-1.03	-0.06	0.02	-3.21
	Sentence Length (Centred)	0.00	0.04	0.03	0.23	0.02	10.03
	Trial Number	0.00	0.00	-0.90	0.00	0.00	-4.92
	Mood × Truth	-0.02	0.02	-0.72	-0.01	0.02	-0.95
Regression-In Probability	(Intercept)	-0.60	0.31	-1.94	-1.80	0.27	-6.58
	Mood: Subjunctive vs. Indicative	-0.44	0.17	-2.54	-0.08	0.10	-0.81
	Truth: True vs. False	-0.26	0.12	-2.26	-0.06	0.10	-0.61
	Sentence Length (Centred)	-0.08	0.18	-0.46	0.19	0.10	1.83
	Trial Number	-0.01	0.01	-1.39	-0.01	0.01	-1.50
	Mood × Truth	-0.03	0.11	-0.29	-0.11	0.10	-1.09
Regression-Out Probability	(Intercept)	-2.76	0.54	-5.15	-2.12	0.28	-7.60
	Mood: Subjunctive vs. Indicative	1.04	0.61	1.71	0.03	0.10	0.36
	Truth: True vs. False	0.17	0.29	0.61	-0.11	0.10	-1.10
	Sentence Length (Centred)	0.57	0.52	1.10	0.11	0.14	0.79
	Trial Number	-0.02	0.01	-1.40	0.01	0.01	1.14
	Mood × Truth	0.10	0.29	0.34	-0.14	0.10	-1.41

Note: Shaded cells denote statistical significance ($p < .05$).

Table 5.15 Mixed-effects models on L2 local in comprehension

		Critical Region			Spillover Region		
		b	*SE*	*t*/*z*	*b*	*SE*	*t*/*z*
First Fixation Duration	(Intercept)	5.57	0.08	69.37	5.48	0.07	79.19
	Mood: Subjunctive vs. Indicative	0.02	0.05	0.35	0.03	0.04	0.85
	Truth: True vs. False	0.01	0.06	0.12	-0.02	0.04	-0.50
	Proficiency (Centred)	0.13	0.08	1.64	0.07	0.07	1.03
	Residence (in French-Speaking Country)	-0.15	0.08	-1.77	-0.12	0.07	-1.61
	Sentence Length (Centred)	0.02	0.02	0.95	0.01	0.01	0.83
	Trial Number	0.00	0.00	-1.82	0.00	0.00	0.21
	Mood × Truth	-0.09	0.05	-1.76	-0.01	0.04	-0.34
	Mood × Prof.	-0.01	0.05	-0.29	0.01	0.04	0.33
	Truth × Prof.	0.02	0.06	0.29	0.01	0.04	0.20
	Mood × Residence	0.01	0.05	0.13	-0.03	0.04	-0.71
	Truth × Residence	0.02	0.06	0.30	0.03	0.04	0.62
	Proficiency × Residence	-0.14	0.09	-1.47	-0.12	0.08	-1.42
	Mood × Truth × Prof.	-0.08	0.05	-1.62	-0.04	0.04	-0.94
	Mood × Truth × Residence	0.10	0.05	1.83	-0.02	0.04	-0.44
	Mood × Prof. × Residence	0.00	0.06	-0.03	-0.03	0.04	-0.60
	Truth × Prof. × Residence	-0.01	0.07	-0.20	0.02	0.04	0.42
	Mood × Truth × Prof. × Residence	0.10	0.06	1.71	0.02	0.04	0.49
Gaze Duration	(Intercept)	5.60	0.09	63.27	5.91	0.11	55.93
	Mood: Subjunctive vs. Indicative	0.03	0.06	0.60	0.03	0.06	0.51
	Truth: True vs. False	0.00	0.05	-0.05	-0.08	0.05	-1.50
	Proficiency (Centred)	0.14	0.09	1.56	0.11	0.11	1.06
	Residence (in French-Speaking Country)	-0.16	0.09	-1.68	-0.25	0.11	-2.27
	Sentence Length (Centred)	0.02	0.02	1.18	0.22	0.03	7.44
	Trial Number	0.00	0.00	-2.19	0.00	0.00	-0.35
	Mood × Truth	-0.09	0.05	-1.68	0.04	0.05	0.80
	Mood × Prof.	0.00	0.06	-0.06	0.02	0.06	0.39
	Truth × Prof.	0.01	0.06	0.23	-0.03	0.06	-0.52
	Mood × Residence	0.00	0.06	0.08	-0.05	0.06	-0.81
	Truth × Residence	0.04	0.06	0.62	0.06	0.06	0.99
	Proficiency × Residence	-0.16	0.10	-1.53	-0.23	0.12	-1.89
	Mood × Truth × Prof.	-0.08	0.06	-1.49	0.00	0.06	0.04
	Mood × Truth × Residence	0.11	0.06	1.90	-0.07	0.06	-1.08
	Mood × Prof. × Residence	-0.03	0.06	-0.40	-0.07	0.07	-0.93
	Truth × Prof. × Residence	-0.01	0.06	-0.15	0.07	0.07	1.09
	Mood × Truth × Prof. × Residence	0.09	0.06	1.48	-0.01	0.07	-0.08
Go-Past Time	(Intercept)	5.61	0.09	65.01	6.10	0.12	50.59
	Mood: Subjunctive vs. Indicative	0.06	0.06	0.94	0.06	0.05	1.25
	Truth: True vs. False	0.00	0.05	0.04	-0.11	0.05	-2.27
	Proficiency (Centred)	0.10	0.09	1.10	0.12	0.12	0.94
	Residence (in French-Speaking Country)	-0.13	0.09	-1.42	-0.23	0.13	-1.80
	Sentence Length (Centred)	0.01	0.02	0.45	0.29	0.03	9.23
	Trial Number	0.00	0.00	-2.46	0.00	0.00	-2.23
	Mood × Truth	-0.08	0.05	-1.47	0.04	0.05	0.78
	Mood × Prof.	0.07	0.07	1.02	0.05	0.05	1.05
	Truth × Prof.	0.04	0.06	0.71	-0.06	0.05	-1.15
	Mood × Residence	-0.02	0.07	-0.34	-0.04	0.05	-0.82
	Truth × Residence	0.03	0.06	0.44	0.08	0.05	1.52
	Proficiency × Residence	-0.13	0.10	-1.27	-0.27	0.14	-1.94
	Mood × Truth × Prof.	-0.09	0.06	-1.56	0.05	0.05	1.04
	Mood × Truth × Residence	0.09	0.06	1.53	-0.02	0.05	-0.41
	Mood × Prof. × Residence	-0.09	0.07	-1.16	-0.08	0.06	-1.34
	Truth × Prof. × Residence	-0.05	0.06	-0.73	0.15	0.06	2.66
	Mood × Truth × Prof. × Residence	0.09	0.06	1.48	0.01	0.06	0.15
Total Reading Time	(Intercept)	5.83	0.13	45.97	6.43	0.17	36.90
	Mood: Subjunctive vs. Indicative	0.06	0.08	0.71	-0.03	0.06	-0.54
	Truth: True vs. False	-0.03	0.07	-0.43	-0.13	0.06	-2.26
	Proficiency (Centred)	0.06	0.13	0.45	0.03	0.18	0.17

		Critical Region			Spillover Region		
		b	*SE*	*t*/*z*	*b*	*SE*	*t*/*z*
	Residence (in French-Speaking Country)	-0.07	0.13	-0.51	-0.16	0.19	-0.87
	Sentence Length (Centred)	0.02	0.03	0.67	0.29	0.03	8.44
	Trial Number	0.00	0.00	-3.64	0.00	0.00	-5.34
	Mood × Truth	-0.07	0.07	-1.01	0.00	0.06	0.07
	Mood × Prof.	0.01	0.08	0.10	-0.07	0.06	-1.22
	Truth × Prof.	0.02	0.08	0.32	-0.05	0.06	-0.77
	Mood × Residence	0.03	0.08	0.43	0.00	0.06	-0.03
	Truth × Residence	0.05	0.08	0.62	0.07	0.06	1.08
	Proficiency × Residence	-0.14	0.15	-0.91	-0.19	0.21	-0.91
	Mood × Truth × Prof.	-0.08	0.08	-1.02	0.01	0.06	0.12
	Mood × Truth × Residence	0.09	0.08	1.09	0.01	0.06	0.20
	Mood × Prof. × Residence	-0.07	0.09	-0.81	0.00	0.07	0.06
	Truth × Prof. × Residence	-0.02	0.09	-0.26	0.13	0.07	1.83
	Mood × Truth × Prof. × Residence	0.08	0.09	0.93	0.04	0.07	0.54
Regression-In Probability	(Intercept)	-0.53	0.33	-1.60	-0.92	0.29	-3.15
	Mood: Subjunctive vs. Indicative	0.40	0.31	1.30	-0.53	0.26	-2.01
	Truth: True vs. False	-0.07	0.30	-0.22	0.14	0.26	0.53
	Proficiency (Centred)	0.07	0.32	0.22	0.17	0.29	0.60
	Residence (in French-Speaking Country)	-0.03	0.33	-0.09	-0.08	0.30	-0.28
	Sentence Length (Centred)	0.09	0.10	0.91	0.20	0.06	3.07
	Trial Number	0.00	0.00	-0.46	-0.01	0.00	-2.42
	Mood × Truth	-0.21	0.30	-0.70	-0.16	0.26	-0.59
	Mood × Prof.	0.18	0.32	0.57	-0.72	0.29	-2.52
	Truth × Prof.	0.04	0.32	0.12	0.23	0.29	0.82
	Mood × Residence	-0.50	0.33	-1.54	0.37	0.30	1.24
	Truth × Residence	0.13	0.33	0.38	-0.26	0.30	-0.89
	Proficiency × Residence	-0.29	0.36	-0.80	-0.37	0.33	-1.13
	Mood × Truth × Prof.	0.01	0.32	0.02	-0.25	0.29	-0.86
	Mood × Truth × Residence	0.39	0.33	1.19	0.35	0.30	1.20
	Mood × Prof. × Residence	-0.58	0.36	-1.60	0.58	0.33	1.76
	Truth × Prof. × Residence	0.16	0.36	0.43	-0.22	0.33	-0.66
	Mood × Truth × Prof. × Residence	0.23	0.36	0.63	0.46	0.33	1.40
Regression-Out Probability	(Intercept)	-2.50	0.72	-3.49	-1.56	0.31	-4.99
	Mood: Subjunctive vs. Indicative	0.11	0.68	0.16	0.36	0.28	1.27
	Truth: True vs. False	0.72	0.66	1.10	-0.42	0.28	-1.49
	Proficiency (Centred)	0.39	0.74	0.53	-0.11	0.30	-0.38
	Residence (in French-Speaking Country)	-0.32	0.75	-0.44	0.44	0.31	1.44
	Sentence Length (Centred)	-0.12	0.19	-0.65	0.14	0.07	2.08
	Trial Number	-0.01	0.01	-0.71	0.00	0.00	-0.19
	Mood × Truth	0.21	0.66	0.31	-0.06	0.28	-0.23
	Mood × Prof.	0.38	0.74	0.52	0.27	0.30	0.90
	Truth × Prof.	0.86	0.74	1.17	-0.50	0.30	-1.66
	Mood × Residence	-0.50	0.75	-0.67	-0.25	0.31	-0.81
	Truth × Residence	-1.30	0.75	-1.74	0.43	0.31	1.39
	Proficiency × Residence	-0.98	0.83	-1.19	0.12	0.34	0.36
	Mood × Truth × Prof.	-0.15	0.74	-0.20	0.25	0.30	0.83
	Mood × Truth × Residence	-0.58	0.75	-0.78	0.23	0.31	0.75
	Mood × Prof. × Residence	-0.75	0.83	-0.91	-0.32	0.34	-0.95
	Truth × Prof. × Residence	-1.58	0.83	-1.91	0.65	0.34	1.92
	Mood × Truth × Prof. × Residence	-0.07	0.83	-0.08	-0.10	0.34	-0.30

Note: Shaded cells denote statistical significance ($p < 0.05$).

5.8.3 Local Measures

5.8.3.1 L1 Group

To examine the overall effect of mood and its interaction with truth-value on reading behaviour, we calculated the following local measures: first fixation duration, gaze duration, go-past time, total reading time, regression-in probability and regression-out probability. For a more in-depth description of these measures, see Section 2.2.4.1. The means (and standard deviations) and the fixed effect estimates for the local measures in the critical and spillover region are shown in Table 5.12 and Table 5.14, respectively.

In the critical region, we only found a significant main effect of mood for regression-in probability, which was higher for indicative than subjunctive sentences. A significant main effect of truth-value was only observed for regression-in probability, such that it was higher for false than indicative sentences. None of the local measures exhibited a significant interaction between mood and truth-value.

In the spillover region, we did not find a significant main effect of mood for any of the local measures. A significant main effect of truth-value was observed, however, for gaze duration, go-past time and total reading time. In particular, these measures were higher for false than true sentences. None of the local measures exhibited a significant interaction between mood and truth.

5.8.3.2 L2 Group

To examine L2 learners' sensitivity to mood and the influence of truth-value, proficiency and residence abroad on reading behaviour, we calculated the following local measures: first fixation duration, gaze duration, go-past time, total reading time, regression-in probability and regression-out probability. For a more in-depth description of these measures, see Section 2.2.4.1. The means (and standard deviations) and the fixed effect estimates for the local measures in the critical and spillover region are shown in Table 5.13 and Table 5.15, respectively.

In the critical region, we did not find a significant main effect of mood or truth-value, or interaction between mood and truth-value, for any of the local measures. Proficiency and/or residence abroad did not appear to modulate this in any way either.

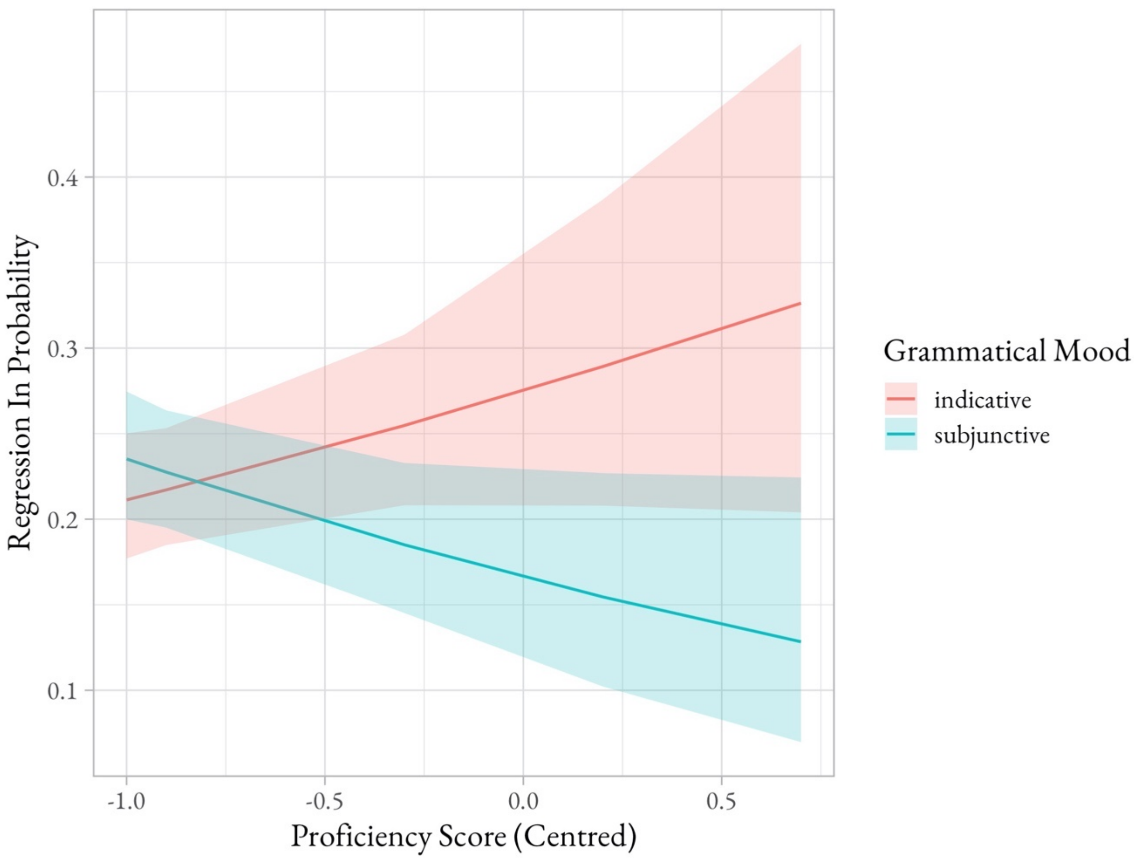


Figure 5.3 Predicted regression-in probability in the spillover region (comprehension)

In the spillover region, there was a significant main effect of mood and a significant interaction between mood and proficiency for regression-in probability, such that regression-in probability was consistently higher for indicative than subjunctive sentences, with proficiency further modulating this effect (see Figure 5.3). A significant main effect of truth-value was observed for both go-past time and total reading time; these reading times were, on average, higher for false than true sentences. For go-past time, there was a significant interaction between truth-value, proficiency and residence abroad for go-past time. Figure 5.4 visualises this interaction. It shows that although false sentences attracted numerically longer go-past times than true sentences among the non-residence abroad group, and that this difference increased as a function of proficiency. In contrast, false sentences received significantly longer go-past times than true sentences at lower levels of proficiency among the residence abroad group, but this pattern reversed at the highest levels of proficiency.

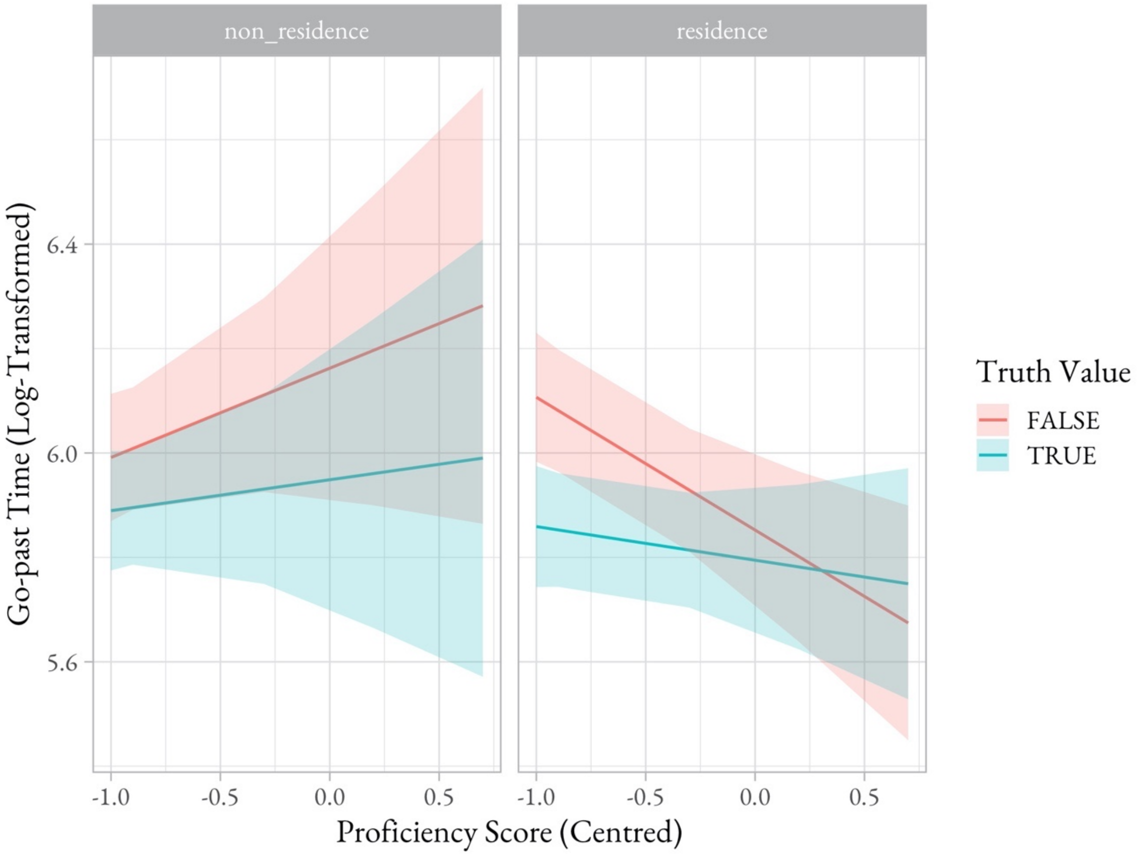


Figure 5.4 Predicted go-past time in the spillover region (comprehension)

Table 5.16 Mixed-effects models for L1 global in judgment

		b	SE	*t*
Mean Fixation Count	(Intercept)	18.39	1.87	9.83
	Mood: Subjunctive vs. Indicative	-0.08	0.26	-0.30
	Truth: True vs. False	0.14	0.29	0.48
	Sentence Length (Centred)	0.78	0.30	2.64
	Trial Number	-0.07	0.01	-5.22
	Mood × Truth	0.38	0.26	1.49
Mean Fixation Duration	(Intercept)	5.43	0.03	202.37
	Mood: Subjunctive vs. Indicative	0.00	0.01	-0.67
	Truth: True vs. False	-0.02	0.01	-3.40
	Sentence Length (Centred)	-0.02	0.01	-2.58
	Trial Number	0.00	0.00	-3.20
	Mood × Truth	0.00	0.01	0.10
Total Sentence Reading Time	(Intercept)	8.67	0.11	81.52
	Mood: Subjunctive vs. Indicative	-0.02	0.01	-1.51
	Truth: True vs. False	-0.01	0.02	-0.73
	Sentence Length (Centred)	0.03	0.02	1.62
	Trial Number	-0.01	0.00	-6.95
	Mood × Truth	0.02	0.01	1.62
Acceptability Rating	(Intercept)	6.18	0.28	22.06
	Mood: Subjunctive vs. Indicative	0.78	0.06	12.15
	Truth: True vs. False	0.04	0.07	0.53
	Sentence Length (Centred)	0.05	0.08	0.67
	Trial Number	0.00	0.00	-1.39
	Mood × Truth	-0.07	0.06	-1.11

Note: Shaded cells denote statistical significance ($p < 0.05$).

Table 5.17 Mixed-effects models for L2 global in judgment

		b	*SE*	*t*
Mean Fixation Count	(Intercept)	15.88	3.30	4.82
	Mood: Subjunctive vs. Indicative	-0.42	0.66	-0.64
	Truth: True vs. False	0.29	0.66	0.44
	Proficiency (Centred)	-5.55	3.50	-1.59
	Residence (in French-Speaking Country)	8.92	3.69	2.42
	Sentence Length (Centred)	0.38	0.19	1.95
	Trial Number	-0.11	0.01	-11.37
	Mood × Truth	-0.18	0.66	-0.27
	Mood × Prof.	-0.19	0.70	-0.27
	Truth × Prof.	1.16	0.70	1.65
	Mood × Residence	0.25	0.74	0.34
	Truth × Residence	-0.97	0.74	-1.30
	Proficiency × Residence	3.55	4.05	0.88
	Mood × Truth × Prof.	0.14	0.70	0.20
	Mood × Truth × Residence	0.19	0.74	0.26
	Mood × Prof. × Residence	-0.08	0.81	-0.10
	Truth × Prof. × Residence	-1.01	0.81	-1.24
	Mood × Truth × Prof. × Residence	0.48	0.81	0.59
Mean Fixation Duration	(Intercept)	5.52	0.13	42.16
	Mood: Subjunctive vs. Indicative	-0.08	0.03	-2.91
	Truth: True vs. False	-0.02	0.03	-0.61
	Proficiency (Centred)	0.02	0.15	0.12
	Residence (in French-Speaking Country)	-0.08	0.14	-0.61
	Sentence Length (Centred)	-0.01	0.01	-0.89
	Trial Number	0.00	0.00	1.29
	Mood × Truth	-0.03	0.03	-1.05
	Mood × Prof.	-0.07	0.03	-2.28
	Truth × Prof.	0.00	0.03	-0.03
	Mood × Residence	0.07	0.03	2.51
	Truth × Residence	0.02	0.03	0.61
	Proficiency × Residence	-0.10	0.16	-0.61
	Mood × Truth × Prof.	-0.02	0.03	-0.76
	Mood × Truth × Residence	0.02	0.03	0.63
	Mood × Prof. × Residence	0.06	0.03	2.04
	Truth × Prof. × Residence	0.01	0.03	0.47
	Mood × Truth × Prof. × Residence	0.01	0.03	0.21
Total Sentence Reading Time	(Intercept)	8.82	0.43	20.44
	Mood: Subjunctive vs. Indicative	-0.11	0.08	-1.46
	Truth: True vs. False	0.12	0.08	1.51
	Proficiency (Centred)	0.06	0.50	0.11
	Residence (in French-Speaking Country)	0.05	0.45	0.12
	Sentence Length (Centred)	0.04	0.02	2.17
	Trial Number	-0.01	0.00	-9.18
	Mood × Truth	0.11	0.08	1.44
	Mood × Prof.	-0.07	0.09	-0.77
	Truth × Prof.	0.18	0.09	2.03
	Mood × Residence	0.09	0.08	1.21
	Truth × Residence	-0.09	0.08	-1.19
	Proficiency × Residence	-0.27	0.52	-0.52
	Mood × Truth × Prof.	0.09	0.09	1.06
	Mood × Truth × Residence	-0.12	0.08	-1.47
	Mood × Prof. × Residence	0.04	0.09	0.39
	Truth × Prof. × Residence	-0.13	0.09	-1.44
	Mood × Truth × Prof. × Residence	-0.12	0.09	-1.29
Acceptability Rating	(Intercept)	4.63	1.36	3.40
	Mood: Subjunctive vs. Indicative	1.78	1.53	1.17
	Truth: True vs. False	-0.01	0.27	-0.03
	Proficiency (Centred)	-1.65	1.56	-1.06

	b	*SE*	*t*
Residence (in French-Speaking Country)	1.39	1.41	0.98
Sentence Length (Centred)	0.00	0.07	-0.04
Trial Number	-0.01	0.00	-5.15
Mood × Truth	-1.28	0.27	-4.77
Mood × Prof.	1.13	1.75	0.65
Truth × Prof.	0.11	0.31	0.35
Mood × Residence	-0.94	1.59	-0.59
Truth × Residence	-0.36	0.28	-1.27
Proficiency × Residence	1.28	1.63	0.79
Mood × Truth × Prof.	-1.31	0.31	-4.26
Mood × Truth × Residence	1.38	0.28	4.94
Mood × Prof. × Residence	-1.01	1.84	-0.55
Truth × Prof. × Residence	-0.71	0.32	-2.19
Mood × Truth × Prof. × Residence	1.49	0.32	4.61

Note: Shaded cells denote statistical significance ($p < 0.05$).

5.9 Results: Study 2b (Judgment)

5.9.1 Data Preparation and Analyses

Data preparation and analyses for Study 2b were identical to those of Study 2a.

5.9.2 Global Measures

5.9.2.1 L1 Group

To examine the overall effect of mood and its interaction with truth-value on reading behaviour, we calculated the following global measures: mean fixation duration, mean fixation count, total sentence reading time and acceptability judgment rating. For each measure, observed means are presented in Table 5.18 and model parameters in Table 5.16.

Table 5.18 Raw means (standard deviations) for L1 global in judgment

1		Total Sentence Reading Time	Mean Fixation Count	Mean Fixation Duration	Acceptability Rating
Indicative	False	5491.69 (2857.74)	15.96 (8.26)	226.88 (35.19)	5.45 (2.04)
	True	5060.03 (2797.2)	14.83 (8.47)	223.76 (36.42)	5.66 (1.9)
Subjunctive	False	5051.53 (2432.68)	15.14 (7.70)	225.22 (35.53)	6.67 (1.07)
	True	5013.74 (2794.29)	15.17 (8.33)	218.86 (31.98)	6.76 (0.82)

Among the global measures, there was only a significant main effect of mood for the acceptability ratings, such that subjunctive sentences were rated as more grammatical than indicative sentences. A significant main effect of truth-value was observed for mean fixation duration, which was longer for false than true sentences. None of the global measures demonstrated a significant interaction between mood and truth.

5.9.2.2 L2 Group

To examine L2 learners' sensitivity to mood and the influence of truth-value, proficiency and residence abroad on reading behaviour, we calculated the following global measures: mean fixation duration, mean fixation count, total sentence reading time and acceptability judgment rating. For each measure, observed means are presented in Table 5.19 and model parameters in Table 5.17.

Table 5.19 Raw means (standard deviations) for L2 global in judgment

		Total Sentence Reading Time	Mean Fixation Count	Mean Fixation Duration	Acceptability Rating
Indicative	False	6839.32 (3113.73)	20.29 (9.40)	242.73 (33.43)	4.58 (2.16)
	True	6458.08 (3140.87)	19.07 (9.09)	240.3 (34.15)	4.67 (2.17)
Subjunctive	False	6832.34 (3193.62)	20.89 (9.45)	239.26 (31.28)	6.46 (1.11)
	True	6048.44 (2799.31)	18.34 (8.44)	235.36 (32.38)	6.42 (1.12)

Among the global measures, we detected a significant main effect of mood, two significant two-way interactions between mood and proficiency and between mood and residence abroad and a significant three-way interaction between mood, proficiency and residence abroad for mean fixation duration. Figure 5.5 visualises this interaction.

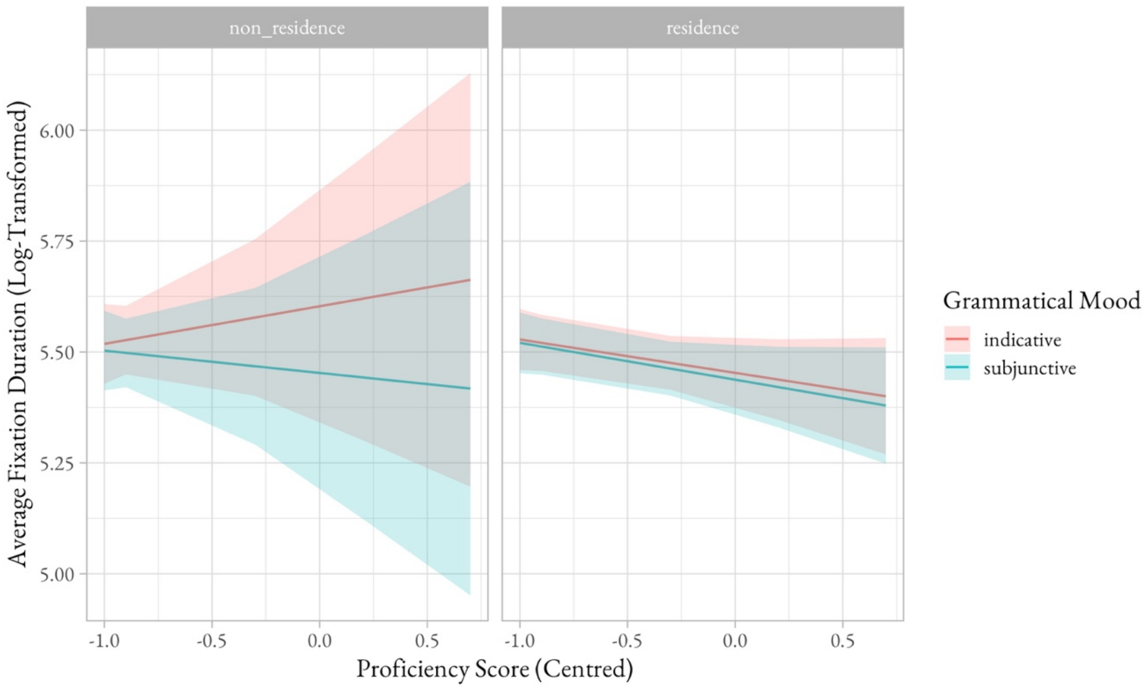


Figure 5.5 Predicted mean fixation duration (judgment)

It shows that although mean fixation duration was numerically longer for indicative than subjunctive sentences, particularly as a function of proficiency among the non-residence abroad group. Among the residence abroad group, however, mean fixation duration was only marginally higher for indicative sentences.

We further found a significant two-way interaction between truth-value and proficiency for total reading time. Figure 5.6 visualises this interaction. It reveals that initially total sentence reading time was numerically longer for false than true sentences. However, as proficiency increased, this trend reversed to the extent that total sentence reading time became numerically longer for true than false sentences.

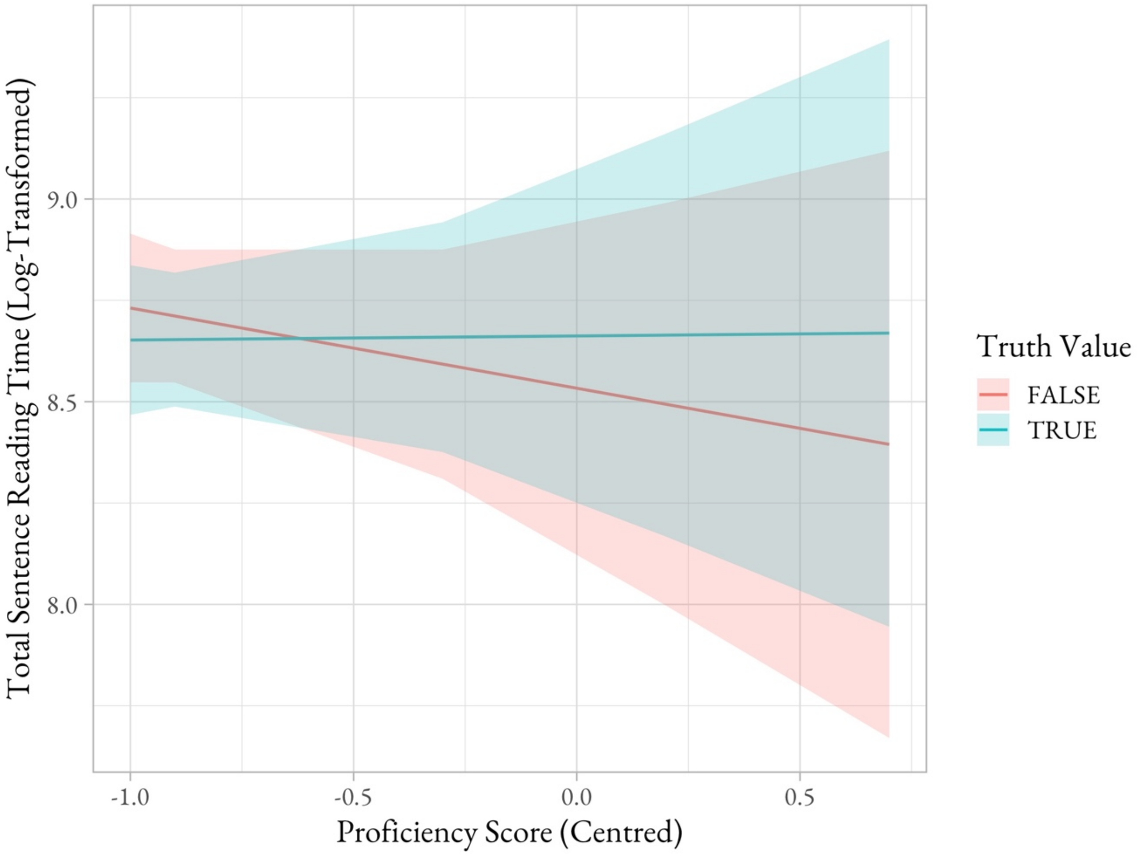


Figure 5.6 Predicted total sentence reading time (judgment)

Finally, we detected a significant two-way interaction between mood and truth, two significant three-way interactions between mood, truth and proficiency and between mood, truth and residence abroad, and a significant four-way interaction between mood, truth, proficiency and residence abroad for acceptability judgment ratings. Figure 5.7 explores this interaction. It reveals that subjunctive sentences consistently attracted higher ratings, but that proficiency and residence abroad significantly modulated this pattern, particularly among false sentences in the non-residence abroad group.

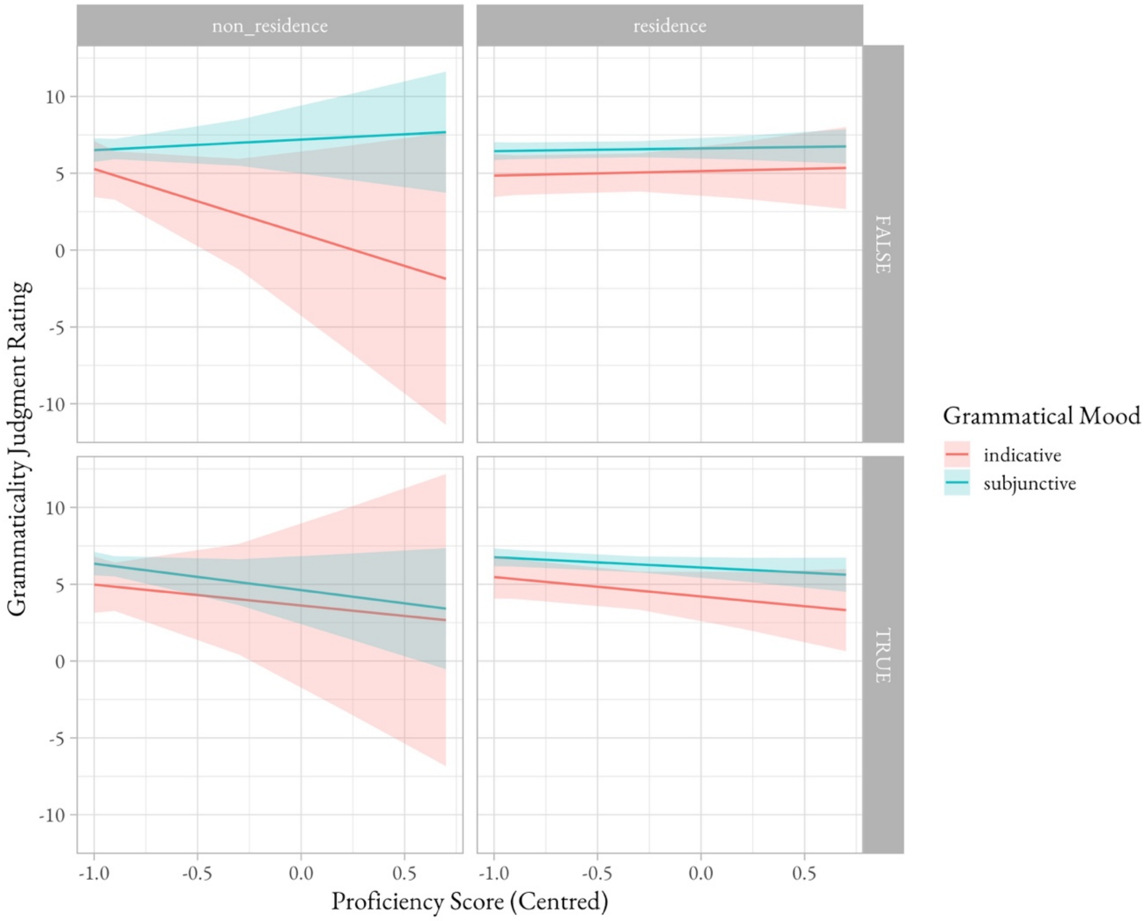


Figure 5.7 Predicted acceptability judgment ratings

.

Table 5.20 Raw means (standard deviations) for L1 local in judgment

			First Fixation Duration	Gaze Duration	Go-Past Time	Total Reading Time	Skipping Probability	Regression-In Probability	Regression-Out Probability
Critical Region	Indicative	False	230.1 (119.32)	240.76 (137.62)	250.05 (139.14)	388.25 (364.86)	0.85 (0.35)	0.46 (0.5)	0.06 (0.25)
		True	241.65 (131.33)	258.11 (182.09)	264.13 (187.28)	370.67 (316.33)	0.91 (0.29)	0.54 (0.5)	0.06 (0.23)
	Subjunctive	False	219.28 (88.99)	225.82 (96.78)	228.5 (97.99)	368.25 (345.27)	0.71 (0.46)	0.27 (0.44)	0.04 (0.19)
		True	224.38 (83.34)	230.18 (89.60)	250.81 (143.20)	386.03 (381.64)	0.78 (0.41)	0.28 (0.45)	0.09 (0.28)
Spillover Region	Indicative	False	224.58 (71.05)	326.31 (219.77)	390.6 (269.33)	592.27 (487.52)	0.48 (0.50)	0.18 (0.39)	0.16 (0.37)
		True	222.1 (78.21)	293.01 (171.19)	357.96 (269.5)	607.27 (634.57)	0.45 (0.50)	0.24 (0.43)	0.17 (0.37)
	Subjunctive	False	213.45 (69.06)	270.85 (138.59)	305.06 (150.8)	470.72 (350.66)	0.37 (0.48)	0.15 (0.36)	0.14 (0.34)
		True	210.39 (65.96)	264.9 (141.03)	306.62 (177.89)	455.98 (375.57)	0.43 (0.50)	0.13 (0.34)	0.17 (0.37)

Table 5.21 Raw means (standard deviations) for L2 local in judgment

			First Fixation Duration	Gaze Duration	Go-Past Time	Total Reading Time	Skipping Probability	Regression-In Probability	Regression-Out Probability
Critical Region	Indicative	False	255.93 (111.18)	269.46 (121.66)	284.02 (144.67)	420.78 (253.22)	0.78 (0.42)	0.37 (0.49)	0.1 (0.3)
		True	273.68 (114.48)	288.09 (126.41)	309.36 (158.53)	453.98 (288.13)	0.76 (0.43)	0.39 (0.49)	0.09 (0.29)
	Subjunctive	False	250.25 (101.52)	261.7 (113.97)	282.33 (132.1)	483.71 (345.01)	0.68 (0.47)	0.36 (0.48)	0.08 (0.28)
		True	243.68 (95.86)	253.68 (108.09)	266.25 (115.67)	403.98 (265.76)	0.63 (0.48)	0.31 (0.46)	0.1 (0.3)
Spillover Region	Indicative	False	247.07 (88.18)	360.54 (224.16)	451.76 (303.61)	732.14 (508.57)	0.36 (0.48)	0.23 (0.42)	0.21 (0.41)
		True	239.04 (93.17)	378.09 (316.4)	466.23 (376.78)	773.4 (623.95)	0.40 (0.49)	0.28 (0.45)	0.2 (0.4)
	Subjunctive	False	245.51 (84.59)	365.96 (214.04)	458.36 (286.33)	776.37 (630.66)	0.36 (0.48)	0.23 (0.42)	0.16 (0.37)
		True	240.03 (98.35)	350.52 (242.54)	449.09 (335.63)	725.41 (556.8)	0.39 (0.49)	0.26 (0.44)	0.2 (0.40)

Table 5.22 Mixed-effects models for L1 local in judgment

		Critical Region			Spillover Region		
		b	*SE*	*t*/*z*	*b*	*SE*	*t*/*z*
First Fixation Duration	(Intercept)	5.31	0.05	111.81	5.35	0.03	178.30
	Mood: Subjunctive vs. Indicative	-0.02	0.02	-0.81	-0.02	0.01	-2.13
	Truth: True vs. False	0.02	0.02	1.08	-0.01	0.01	-0.83
	Sentence Length (Centred)	-0.02	0.02	-0.80	-0.01	0.01	-0.45
	Trial Number	0.00	0.00	0.51	0.00	0.00	-0.74
	Mood × Truth	0.00	0.02	-0.21	0.00	0.01	0.25
Gaze Duration	(Intercept)	5.32	0.05	104.46	5.53	0.04	123.08
	Mood: Subjunctive vs. Indicative	-0.03	0.02	-1.09	-0.06	0.01	-4.14
	Truth: True vs. False	0.02	0.02	1.05	-0.03	0.01	-2.06
	Sentence Length (Centred)	-0.02	0.02	-0.99	0.14	0.02	6.41
	Trial Number	0.00	0.00	0.78	0.00	0.00	0.22
	Mood × Truth	0.00	0.02	-0.27	0.02	0.01	1.39
Go-Past Time	(Intercept)	5.32	0.05	100.40	5.67	0.05	125.26
	Mood: Subjunctive vs. Indicative	-0.02	0.03	-0.74	-0.08	0.01	-5.93
	Truth: True vs. False	0.02	0.02	1.28	-0.03	0.01	-2.33
	Sentence Length (Centred)	-0.02	0.03	-0.67	0.22	0.01	15.89
	Trial Number	0.00	0.00	1.30	0.00	0.00	0.35
	Mood × Truth	0.01	0.02	0.55	0.02	0.01	1.35
Total Reading Time	(Intercept)	5.60	0.09	62.59	6.05	0.08	73.36
	Mood: Subjunctive vs. Indicative	0.02	0.04	0.42	-0.10	0.02	-6.14
	Truth: True vs. False	-0.01	0.02	-0.23	-0.03	0.02	-1.43
	Sentence Length (Centred)	-0.01	0.05	-0.21	0.25	0.02	10.86
	Trial Number	0.00	0.00	-0.48	0.00	0.00	-2.02
	Mood × Truth	0.00	0.02	-0.06	-0.01	0.02	-0.49
Regression-In Probability	(Intercept)	-0.51	0.33	-1.55	-1.57	0.27	-5.92
	Mood: Subjunctive vs. Indicative	-0.66	0.15	-4.48	-0.28	0.09	-3.00
	Truth: True vs. False	0.13	0.10	1.20	0.03	0.09	0.33
	Sentence Length (Centred)	-0.19	0.14	-1.37	0.30	0.10	3.00
	Trial Number	-0.01	0.01	-1.78	-0.01	0.01	-1.48
	Mood × Truth	-0.08	0.10	-0.78	-0.15	0.09	-1.56
Regression-Out Probability	(Intercept)	-3.62	0.54	-6.71	-2.30	0.29	-7.89
	Mood: Subjunctive vs. Indicative	0.18	0.30	0.61	-0.07	0.10	-0.71
	Truth: True vs. False	0.21	0.20	1.02	0.07	0.10	0.74
	Sentence Length (Centred)	0.25	0.27	0.92	0.30	0.15	2.03
	Trial Number	0.02	0.01	1.39	0.01	0.01	1.57
	Mood × Truth	0.27	0.20	1.33	0.04	0.10	0.47

Note: Shaded cells denote statistical significance ($p < .05$).

Table 5.23 Mixed-effects models for L2 local in judgment

		Critical Region			Spillover Region		
		b	*SE*	*t/z*	*b*	*SE*	*t/z*
First Fixation Duration	(Intercept)	5.51	0.10	54.73	5.39	0.07	77.19
	Mood: Subjunctive vs. Indicative	-0.06	0.05	-1.25	-0.02	0.04	-0.51
	Truth: True vs. False	0.02	0.05	0.44	-0.04	0.03	-1.25
	Proficiency (Centred)	0.04	0.11	0.40	0.01	0.07	0.07
	Residence (in French-Speaking Country)	-0.02	0.11	-0.19	0.00	0.08	-0.05
	Sentence Length (Centred)	-0.04	0.02	-2.29	0.00	0.01	-0.40
	Trial Number	0.00	0.00	-1.88	0.00	0.00	2.17
	Mood × Truth	-0.13	0.05	-2.91	-0.02	0.03	-0.50
	Mood × Prof.	-0.02	0.05	-0.41	-0.02	0.04	-0.46
	Truth × Prof.	0.03	0.05	0.61	-0.02	0.04	-0.61
	Mood × Residence	-0.03	0.05	-0.63	0.03	0.04	0.62
	Truth × Residence	0.02	0.05	0.45	0.03	0.04	0.89
	Proficiency × Residence	-0.09	0.12	-0.73	-0.02	0.08	-0.19
	Mood × Truth × Prof.	-0.10	0.05	-2.01	0.00	0.04	0.04
	Mood × Truth × Residence	0.12	0.05	2.40	0.03	0.04	0.85
	Mood × Prof. × Residence	-0.02	0.06	-0.41	0.03	0.05	0.66
	Truth × Prof. × Residence	-0.01	0.06	-0.11	0.03	0.04	0.77
	Mood × Truth × Prof. × Residence	0.09	0.06	1.65	0.00	0.04	-0.06
Gaze Duration	(Intercept)	5.55	0.11	48.63	5.72	0.11	49.99
	Mood: Subjunctive vs. Indicative	-0.05	0.05	-1.12	-0.01	0.06	-0.21
	Truth: True vs. False	0.02	0.05	0.42	-0.07	0.05	-1.41
	Proficiency (Centred)	0.06	0.12	0.51	0.07	0.12	0.56
	Residence (in French-Speaking Country)	-0.04	0.13	-0.32	-0.12	0.12	-0.94
	Sentence Length (Centred)	-0.02	0.02	-1.07	0.16	0.02	7.45
	Trial Number	0.00	0.00	-1.74	0.00	0.00	2.21
	Mood × Truth	-0.14	0.05	-2.94	-0.05	0.05	-1.02
	Mood × Prof.	-0.04	0.05	-0.72	-0.03	0.06	-0.44
	Truth × Prof.	0.03	0.05	0.56	-0.07	0.06	-1.26
	Mood × Residence	-0.03	0.05	-0.56	0.02	0.07	0.37
	Truth × Residence	0.01	0.05	0.26	0.03	0.06	0.57
	Proficiency × Residence	-0.12	0.14	-0.86	-0.13	0.14	-0.99
	Mood × Truth × Prof.	-0.10	0.05	-1.99	-0.02	0.06	-0.35
	Mood × Truth × Residence	0.14	0.05	2.55	0.06	0.06	1.01
	Mood × Prof. × Residence	0.00	0.06	-0.08	0.08	0.07	1.13
	Truth × Prof. × Residence	-0.01	0.06	-0.25	0.07	0.07	1.13
	Mood × Truth × Prof. × Residence	0.10	0.06	1.69	0.02	0.06	0.25
Go-Past Time	(Intercept)	5.59	0.12	46.93	5.87	0.13	46.09
	Mood: Subjunctive vs. Indicative	-0.06	0.05	-1.10	-0.06	0.06	-1.10
	Truth: True vs. False	0.01	0.05	0.15	-0.05	0.05	-1.00
	Proficiency (Centred)	0.08	0.13	0.67	0.02	0.13	0.17
	Residence (in French-Speaking Country)	-0.05	0.13	-0.41	-0.02	0.14	-0.16
	Sentence Length (Centred)	0.00	0.02	-0.10	0.23	0.03	7.66
	Trial Number	0.00	0.00	-0.87	0.00	0.00	1.30
	Mood × Truth	-0.11	0.05	-2.30	-0.04	0.05	-0.83
	Mood × Prof.	-0.04	0.05	-0.70	-0.07	0.06	-1.08
	Truth × Prof.	0.01	0.05	0.24	-0.02	0.05	-0.36
	Mood × Residence	0.01	0.06	0.15	0.06	0.07	0.91
	Truth × Residence	0.01	0.06	0.20	0.03	0.06	0.56
	Proficiency × Residence	-0.15	0.14	-1.02	-0.14	0.15	-0.93
	Mood × Truth × Prof.	-0.09	0.05	-1.61	-0.03	0.05	-0.47
	Mood × Truth × Residence	0.08	0.06	1.51	0.04	0.06	0.70
	Mood × Prof. × Residence	0.01	0.06	0.08	0.05	0.07	0.65
	Truth × Prof. × Residence	-0.01	0.06	-0.14	0.04	0.06	0.65
	Mood × Truth × Prof. × Residence	0.07	0.06	1.18	0.03	0.06	0.42
Total Reading Time	(Intercept)	5.76	0.17	33.36	6.12	0.19	31.93
	Mood: Subjunctive vs. Indicative	-0.06	0.08	-0.74	-0.11	0.08	-1.31
	Truth: True vs. False	-0.10	0.07	-1.45	-0.03	0.06	-0.49
	Proficiency (Centred)	-0.04	0.18	-0.24	-0.28	0.20	-1.39

		Critical Region			Spillover Region		
		b	*SE*	*t/z*	*b*	*SE*	*t/z*
	Residence (in French-Speaking Country)	0.15	0.19	0.79	0.37	0.21	1.76
	Sentence Length (Centred)	0.07	0.03	2.60	0.25	0.03	8.24
	Trial Number	0.00	0.00	-2.14	0.00	0.00	-3.87
	Mood × Truth	-0.04	0.07	-0.67	-0.02	0.06	-0.27
	Mood × Prof.	-0.12	0.08	-1.45	-0.10	0.09	-1.15
	Truth × Prof.	-0.08	0.07	-1.19	0.03	0.06	0.43
	Mood × Residence	0.08	0.09	0.94	0.10	0.09	1.11
	Truth × Residence	0.09	0.07	1.19	0.04	0.06	0.64
	Proficiency × Residence	-0.10	0.21	-0.50	0.15	0.23	0.66
	Mood × Truth × Prof.	-0.01	0.07	-0.16	0.00	0.06	-0.02
	Mood × Truth × Residence	0.00	0.07	-0.05	0.03	0.06	0.46
	Mood × Prof. × Residence	0.08	0.10	0.87	0.08	0.10	0.76
	Truth × Prof. × Residence	0.10	0.08	1.28	0.02	0.07	0.30
	Mood × Truth × Prof. × Residence	0.03	0.08	0.39	0.05	0.07	0.72
Regression-In Probability	(Intercept)	-1.19	0.34	-3.54	-1.68	0.30	-5.52
	Mood: Subjunctive vs. Indicative	-0.58	0.32	-1.86	0.16	0.28	0.56
	Truth: True vs. False	-0.17	0.31	-0.53	-0.21	0.28	-0.74
	Proficiency (Centred)	-0.46	0.33	-1.42	-0.76	0.28	-2.71
	Residence (in French-Speaking Country)	0.63	0.34	1.84	0.71	0.31	2.31
	Sentence Length (Centred)	0.07	0.09	0.76	0.23	0.06	3.67
	Trial Number	-0.01	0.00	-1.41	-0.01	0.00	-2.09
	Mood × Truth	0.63	0.31	2.03	-0.12	0.28	-0.42
	Mood × Prof.	-0.58	0.33	-1.77	0.26	0.28	0.92
	Truth × Prof.	-0.26	0.33	-0.80	-0.30	0.28	-1.05
	Mood × Residence	0.44	0.34	1.28	-0.05	0.31	-0.18
	Truth × Residence	0.15	0.34	0.45	0.52	0.31	1.70
	Proficiency × Residence	-0.12	0.37	-0.34	0.47	0.32	1.44
	Mood × Truth × Prof.	0.64	0.33	1.95	-0.31	0.28	-1.11
	Mood × Truth × Residence	-0.79	0.34	-2.32	-0.10	0.31	-0.31
	Mood × Prof. × Residence	0.43	0.37	1.16	-0.14	0.32	-0.43
	Truth × Prof. × Residence	0.38	0.37	1.04	0.53	0.32	1.63
	Mood × Truth × Prof. × Residence	-0.64	0.37	-1.74	0.23	0.32	0.70
Regression-Out Probability	(Intercept)	-2.30	0.52	-4.42	-0.74	0.32	-2.35
	Mood: Subjunctive vs. Indicative	0.21	0.48	0.45	-0.57	0.28	-2.00
	Truth: True vs. False	0.06	0.47	0.12	0.23	0.28	0.81
	Proficiency (Centred)	0.73	0.57	1.27	0.48	0.31	1.55
	Residence (in French-Speaking Country)	-0.18	0.52	-0.35	-0.18	0.32	-0.58
	Sentence Length (Centred)	0.11	0.15	0.74	0.12	0.07	1.76
	Trial Number	0.01	0.01	0.95	-0.01	0.00	-3.57
	Mood × Truth	-0.38	0.47	-0.80	0.47	0.28	1.64
	Mood × Prof.	0.50	0.57	0.87	-0.44	0.31	-1.43
	Truth × Prof.	-0.17	0.57	-0.30	0.05	0.31	0.15
	Mood × Residence	0.07	0.52	0.14	0.45	0.32	1.41
	Truth × Residence	-0.21	0.52	-0.40	-0.17	0.32	-0.53
	Proficiency × Residence	-0.97	0.63	-1.54	-0.63	0.35	-1.80
	Mood × Truth × Prof.	-0.88	0.57	-1.54	0.36	0.31	1.15
	Mood × Truth × Residence	0.45	0.52	0.87	-0.30	0.32	-0.95
	Mood × Prof. × Residence	-0.33	0.63	-0.52	0.29	0.35	0.82
	Truth × Prof. × Residence	-0.05	0.63	-0.08	0.13	0.35	0.36
	Mood × Truth × Prof. × Residence	1.00	0.63	1.59	-0.19	0.35	-0.54

Note: Shaded cells denote statistical significance ($p < .05$).

5.9.3 Local Measures

5.9.3.1 L1 Group

To examine the overall effect of mood and its interaction with truth-value on reading behaviour, we calculated the following local measures: first fixation duration, gaze duration, go-past time, total reading time, regression-in probability and regression-out probability. For a more in-depth description of these measures, see Section 2.2.4.1. The means (and standard deviations) and the fixed effect estimates for the local measures in the critical and spillover region are shown in Table 5.20 and Table 5.22, respectively.

In the critical region, there was a significant main effect of mood for regression-in probability, such that these probabilities were higher for indicative than subjunctive sentences.

In the spillover region, a significant main effect of mood was found for first fixation duration, gaze duration, go-past time, total reading time and regression-in probability, with longer reading times (or higher probabilities) for indicative than subjunctive sentences. There was also a significant main effect of truth-value for gaze duration and go-past time, such that these reading measures were higher with false than true sentences.

5.9.3.2 L2 Group

To examine L2 learners' sensitivity to mood and the influence of truth-value, proficiency and residence abroad on reading behaviour, we calculated the following local measures: first fixation duration, gaze duration, go-past time, total reading time, regression-in probability and regression-out probability. For a more in-depth description of these measures, see Section 2.2.4.1. The means (and standard deviations) and the fixed effect estimates for the local measures in the critical and spillover region are shown in Table 5.21 and Table 5.23, respectively.

In the critical region, we did not find a significant main effect of mood or truth-value for any of the local measures. There was, however, a significant interaction between mood and truth-value for first fixation duration, gaze duration and go-past time. In particular, these reading times were longer for Indicative-True sentences (than Indicative-False, Subjunctive-False and Subjunctive-True). In addition, a significant two-way interaction between mood, truth-value and residence abroad was detected for both first fixation duration and gaze duration. Reading times patterned as follows (from longest to shortest): Indicative-True, Subjunctive-False, Indicative-False and then Subjunctive-True for non-residence abroad learners and Indicative-True, Indicative-False, Subjunctive-True and Subjunctive-False for residence abroad learners.

In the spillover region, we did not observe a significant main effect of mood or truth-value, or a significant interaction between mood and truth-value, for any of the local measures.

5.10 Discussion

In Study 1, we found strong evidence to suggest that native speakers of French were sensitive to mood–modality mismatches in obligatory contexts. However, sensitivity among second language learners was limited and highly constrained by factors such as the lexical-semantic properties, proficiency and residence abroad. Unlike Study 1, the current study only detected partial sensitivity to the discourse-pragmatic constraints governing mood use in French among both the control group and the L2 group. This sensitivity did not appear to be modulated by either proficiency or residence abroad among the L2 learners. As such, the current section seeks to discuss both native speakers' and L2 learners' processing of mood in polarity contexts.

EXPERIMENTAL PREDICTIONS

Recall the initial hypotheses concerning the underlying causes of differences between the native speakers and second language learners of French:

- Hypothesis 1: L2 learners should be more sensitive to the discourse-pragmatic constraints in the offline than the online data, but this sensitivity is not guaranteed.
- Hypothesis 2: As proficiency increases, L2 learners should exhibit higher sensitivity to the discourse-pragmatic constraints.
- Hypothesis 3: Exposure to (naturalistic) input, through residence abroad, should improve sensitivity to discourse-pragmatic constraints.

L1 DATA

As stated in the introduction, there was strong evidence to suggest that native speakers of French were able to detect mood–modality mismatches in obligatory contexts (Study 1). However, this was not always the case in polarity contexts. In fact, the native speakers in Study 2 did not reveal any sensitivity to discourse-pragmatic constraints and exhibited a stronger preference for the subjunctive over the indicative in polarity contexts, particularly in the judgment task. We found that the following (reading) measures were significantly higher for indicative than subjunctive sentences in the judgment task: acceptability judgment ratings, regression-in probability in the critical region and first fixation duration, gaze duration, go-past time, total reading time and regression-in probability in the spillover region. This was also the case for regression-in probability in the critical region of the comprehension task. The fact that significant effects were found in both early and late measures reflects a clear ability not only to detect morphosyntactic violations, but also to perform reanalysis following integration difficulties (Rayner *et al.*, 2004). A notable exception, however, was mean fixation duration in the comprehension task, which was longer for subjunctive than indicative sentences, suggesting that the subjunctive was unexpected in this case. Despite this anomaly, the data largely indicate that native speakers largely considered the subjunctive to be the form *par excellence* under negated epistemics in French, particularly when reading for judgment. This is to be expected considering that two out of the three grammatical conditions required the subjunctive.

It is interesting to further examine the type of sensitivity effects observed in the dataset, at least in terms of sensitivity to the ungrammaticality of the indicative regardless of discourse-pragmatic constraints. For example, most sensitivity effects were either detected in the spillover region of the judgment task in the following three measures: first fixation duration, gaze duration, go-past time, total reading time and regression-in probability. These measures have been shown to reflect sensitivity to lexical properties (Rayner, 2009), syntactic and semantic anomalies (Rayner *et al.*, 2004), lexical and integration difficulties leading to regression to earlier points (Rayner *et al.*, 2004), a late effect on processing (Liversedge, Paterson and Pickering, 1998) and higher-level (e.g., discourse) processing (Rayner and Pollatsek, 1989), respectively. This means that the experimental manipulation disrupted reading not only at the early stages of processing, but also at the later stages when native speakers are required to read for judgment rather than for comprehension. Since these findings suggest that the secondary task may influence reading patterns, we leave any further discussion to Study 3.

The presence of sensitivity effects in the spillover region and their relative absence in the critical region suggests that sensitivity may be delayed in L1 processing. There are several possible reasons why this might be the case. One possible explanation relates to the cognitive burden that the computation of counterfactual thought places on the parser. Previous research has shown that this can be cognitively demanding for many native speakers since it requires the reader to entertain possible words that they know to be false (Ferguson, Scheepers and Sanford, 2010; Nieuwland and Martin, 2012; Nieuwland, 2013; Ferguson and Cane, 2015; Kulakova and Nieuwland, 2016a).

In fact, one study observed that pragmatic skill (i.e., "the ability to apply knowledge of the social-communicative use of language in daily life") predicted online counterfactual comprehension (Kulakova and Nieuwland, 2016a, p. 814). In particular, it found that "individuals who are better at understanding the communicative intentions of others are more likely to reduce knowledge-based expectations in counterfactuals" (p.814). Although our study did not measure pragmatic skill, it is plausible that this level of individual difference may have contributed, at least in part, to the variability in L1 behaviour and explain why evidence of sensitivity effects was not always as robust as expected. It is clear from Figure 5.8 that although the vast majority of native speakers rated subjunctive sentences as grammatical, there was a distinct level of variation in the ratings attributed to indicative sentences, regardless of the discourse-pragmatic constraints.

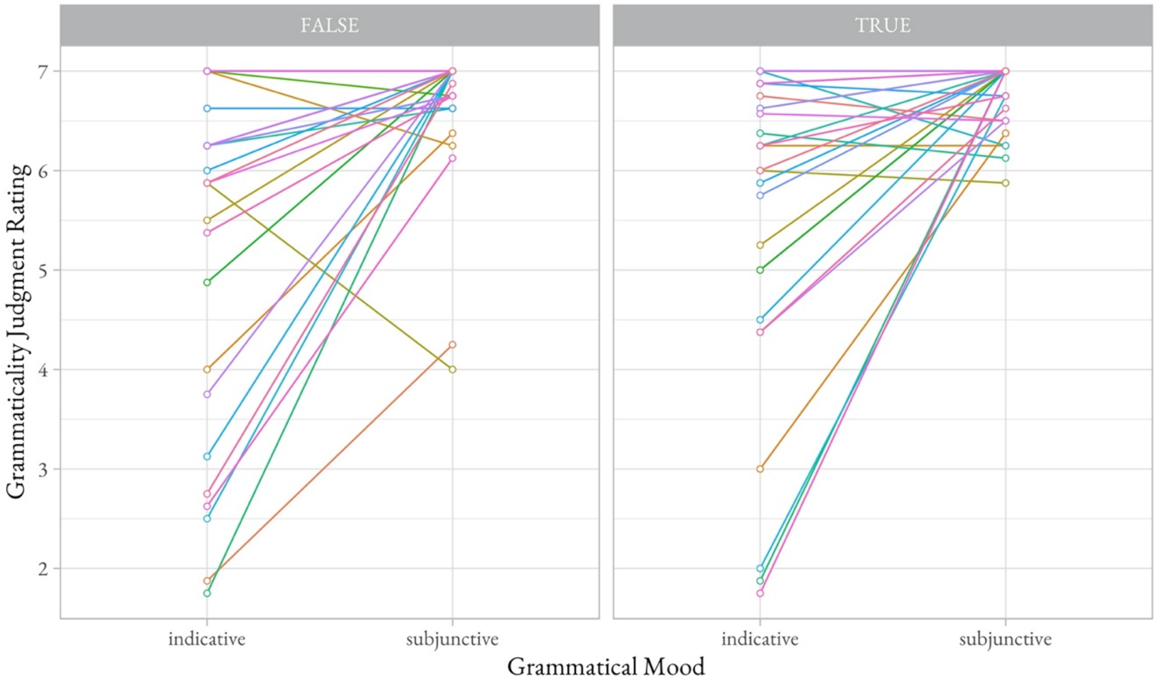


Figure 5.8 Individual L1 participants' grammaticality judgment ratings

An alternative explanation for sensitivity effects in the spillover region relates to the lexical properties of the critical region, in particular its length. The descriptive statistics in Sections 5.8 and 5.9 revealed that the likelihood of a participant skipping the critical region was consistently high across conditions and tasks. This was to be expected given the widely reported direct correlation between skipping probability and the length of the interest area (Balota, Pollatsek and Rayner, 1985; Schustack, Ehrlich and Rayner, 1987; Altarriba *et al.*, 1996; Rayner and Well, 1996). High skipping rates consequently decrease the number of fixations within the interest area and thus limit the number of datapoints available to compute inferential analyses of reading measures. This then reduces statistical power and minimises the likelihood of detecting significant effects.

Finally, it is clear from the findings presented in this study that native speakers did not exhibit any sensitivity to the discourse-pragmatic constraints, only to the ungrammaticality of the indicative. As we have already discussed, one possible explanation for this relates to the cognitive burden of processing counterfactual thought. However, there are several other possible explanations for the lack of consistency in the dataset, including the status of the subjunctive in the French language.

The status of the subjunctive in French has been largely debated over the years. Some scholars have argued that the subjunctive no longer conveys semantic content (Harris, 1978; Mailhac, 2000; Ayres-Bennett and Carruthers, 2001). In fact, some even suggest that the subjunctive is simply a grammatical form preserved by literary traditions and grammatical purism (Levitt, 1967), a marker of formality, almost exclusively restricted to written language (Müller and Elsass, 1985; O'Connor DiVito, 1997; Lenoble-Pinson and Grevisse, 2009; Jeppesen Kragh, 2010). Furthermore, Ayres-Bennett and Carruthers (2001) argue that in spoken French, most French speakers will use the indicative over the subjunctive when the choice presents itself in spoken language. It is therefore plausible that the lack of robustness in the dataset reflects the inter-speaker variation in subjunctive use that has been widely attested in the literature.

L2 DATA

Let us now turn to the L2 data. As we have discussed in the previous sub-section, there was no evidence of sensitivity to discourse-pragmatic constraints, only to the ungrammaticality of the indicative among the control group. We did not detect any sensitivity to discourse-pragmatic constraints among the L2 group either. However, unlike Study 1, this sensitivity did not appear to be conditioned by proficiency and/or residence abroad.

In our initial hypotheses, we anticipated that the successful acquisition of the polarity subjunctive would depend, to a certain extent, on whether the L2 learners had already acquired the relevant properties associated with the obligatory subjunctive. However, evidence of successful acquisition was not always robust in Study 1. Although L2 learners demonstrated consistent offline knowledge of the subjunctive, there was only partial evidence of sensitivity to mood–modality mismatches in the online data. It was therefore argued that the L2 learners did have knowledge of the relevant properties, but that transfer effects often arose in real-time processing due to limited (domain-general) cognitive resources.[35] See Section 4.9 for further discussion of this argument. With this in mind, we chose not to exclude L2 participants based on their reduced sensitivity to mood–modality mismatches during real-time processing in the current study.

[35] Second language processing can be cognitively demanding due to limited domain-general resources (e.g., McDonald, 2006; Hopp, 2010; Sorace, 2011; Dekydtspotter and Renaud, 2014). This may, in turn, lead to L1 transfer effects that do not necessarily surface in offline knowledge. For example, it is possible that the linguistic representations associated with the subjunctive have different activation levels (Dekydtspotter and Renaud, 2014). Because of L1/L2 similarities, desire and directive verbs are likely to have higher activation levels than emotive-factive verbs, and are, as a result, more easily retrieved by the parser.

Although the L1 speakers exhibited a distinct preference for the subjunctive under negated epistemics, this was not the case among the L2 group. In fact, we only observed one measure where L2 speakers demonstrated a preference for one mood over another. This was for the subjunctive with mean fixation duration in the judgment task. This suggests that while L2 speakers may exhibit a marginal (target-like) preference for the subjunctive, this was not always consistent and that L2 speakers displayed a certain level of indeterminacy with respect to verbal mood.

The lack of sensitivity to the discourse-pragmatic constraints and the indeterminacy between the indicative and the subjunctive among the L2 speakers are not surprising at all when you consider the absence of sensitivity to the discourse-pragmatic constraints among the control group. More specifically, it suggests that evidence of the discourse-pragmatic constraints is not always readily available to L2 learners, at least not at the level required for full acquisition.[36] As Slabakova (2015) points out, successful acquisition is often contingent on the availability of abundant, unambiguous input. This is evidently not the case for mood use in polarity contexts. It is therefore not surprising that L2 sensitivity was often limited at best, especially when faced with such reduced and ambiguous input.

Initially, we predicted that increased exposure to the target language through extended residence abroad and proficiency would modulate, or perhaps even determine, sensitivity among L2 learners. This hypothesis, however, turned out to be incorrect. As we briefly mentioned in the introduction, the L2 dataset did not provide conclusive evidence to suggest that proficiency and/or residence abroad played a contributing role. This finding is particularly surprising in light of Study 1, which showed that both residence abroad and proficiency, albeit contradictory in the case of residence abroad, influenced L2 sensitivity to mood–modality mismatches. Further research is needed with a more diverse L2 population to further investigate the influence of proficiency and residence abroad.

Nevertheless, the findings presented above, more specifically the significant level of indeterminacy between the two moods, are largely consistent with the acquisition task predictions formulated within a Feature Reassembly framework (Lardiere, 2009) in Section 5.3.

[36] In fact, it could even be argued that the polarity subjunctive is illustrative of the poverty of the stimulus argument. The poverty of the stimulus refers to the idea that the input is insufficient in accounting for a learner's knowledge of a particular linguistic feature (Lasnik and Lidz, 2016). In the case of the polarity subjunctive, it is almost impossible to evaluate a speaker's (or the matrix subject's) commitment to the truth-value of the embedded proposition from the input alone.

It suggests that L2 learners experience substantial difficulties when required to reconfigure existing feature bundles to include a discourse-based feature, particularly in the context of multiple form-to-meaning mapping. By multiple form-to-meaning mappings, we refer to the fact that the subjunctive can be used to express both commitment and non-commitment to the truth-value of the embedded proposition, whereas the indicative form can only indicate commitment.

An alternative explanation for our findings relates to the role of interface conditions. As we have previously discussed, the polarity subjunctive involves an interface between syntax and discourse-pragmatics, since mood alternation under negated epistemics depends on the speaker's or the matrix subject's commitment to the truth-value of the proposition in embedded clause. The Interface Hypothesis predicts that structures involving external interfaces (e.g., syntax-discourse/pragmatics) may be challenging for L2 learners, either due to "the competition for processing resources created by the need to separate the two languages, or perhaps a specific difficulty in integrating information in real time" (Sorace, 2012, pp. 209–210). Our findings provide evidence in support of the Interface Hypothesis, especially considering that both L1 and L2 learners exhibited reduced sensitivity in polarity contexts, albeit to varying degrees. This is a likely explanation since both groups were English-French bilinguals and as such, were required to separate the two languages, which thus led to an absence of sensitivity due to the cognitive burden placed on the parser. The activation threshold levels of the relevant grammatical features are again likely to come in to play here and be a contributing factor, with higher activation threshold levels among the L1 than the L2 group due to increased contact with the target language.

Finally, as we have previously discussed in relation to the L1 data, previous research has shown that the computation of counterfactual thought in creating possible words has proven to be particularly challenging for individuals who experience difficulties comprehending the communicative intentions of other people (Kulakova and Nieuwland, 2016a). It is plausible that that such difficulties are even more pronounced among the L2 population due to limited cognitive resources resulting from the need to separate the two languages, as hypothesised by limited capacity models of L2 processing (e.g., McDonald, 2006; Hopp, 2010; Sorace, 2011; Dekydtspotter and Renaud, 2014).

To conclude, the current study has revealed a number of interesting findings pertaining to both L1 and L2 sensitivity to the polarity subjunctive. In particular, we found an absence of sensitivity to the discourse-pragmatic constraints governing mood use in French among both the control group and the L2 group. Given that an absence of sensitivity was not limited to just

L2 learners, as was the case in Study 1 (obligatory subjunctives), and was observed among L1 speakers as well, we argued that a number of different factors can explain these findings, including the infrequency of the subjunctive in the input and the cognitive burden that the processing of the polarity subjunctive places on the parser due to its low activation levels in the mental lexicon.

Chapter 6 Study 3: Secondary Task Effects

6.1 Introduction

It has been widely reported that many adult second language (L2) learners face considerable challenges when acquiring inflectional morphology, such as tense, aspect, mood and agreement (Haznedar and Schwartz, 1997; Prévost and White, 2000; Lardiere, 2006; Hopp, 2010; Slabakova, 2013). However, there remains a distinct lack of consensus concerning the exact cause of these difficulties. In recent years, an increasing number of L2 researchers have begun to investigate the extent to which these difficulties reflect differences in sentence processing between the first language (L1) and L2 (Clahsen and Felser, 2006, 2018; Felser and Cunnings, 2012; Clahsen *et al.*, 2013; Jegerski and Sekerina, 2020). More specifically, this research has sought to ascertain whether L2 learners exhibit target-like sensitivity to morphosyntactic violations during online comprehension. Some studies have reported limited sensitivity to morphosyntactic information among L2 learners (Jiang, 2004, 2007; Sato and Felser, 2007; Keating, 2009), whereas others have found that L2 learners are sensitive to morphosyntactic violations and have argued that L1 and L2 processing are not necessarily qualitatively, but rather quantitatively, different (Hopp, 2006; Jackson, 2008; Sagarra and Herschensohn, 2010; Tokowicz and Warren, 2010; Foote, 2011). It has been further suggested that L2 proficiency modulates processing patterns (Frenck-Mestre, 2002; Hopp, 2006; Jackson, 2008; Foote, 2011). Despite this increased interest in L2 sentence processing, little research has considered how task design impacts on L2 processing patterns. In particular, we refer to how the use of secondary tasks, such as comprehension questions and judgment tasks, might influence L2 learners' ability to detect morphosyntactic violations. Existing research has shown that L2 learners are more likely to exhibit target-like processing patterns when the task requires a focus on form rather than function, such as grammaticality judgment tasks and translation (Jackson and Bobb, 2009; Leeser, Brandl and Weissglass, 2011; Lim and Christianson, 2013b, 2015).

The current study aims to further expand this line of enquiry, by investigating whether adult L2 English learners of French of varying degrees of L2 proficiency can detect mood-modality mismatches in two different tasks, namely reading for comprehension versus reading for judgment, during which participants' eye movements were recorded. In the reading for comprehension task, participants answered comprehension questions after 50% of the trials. In the reading for judgment task, participants assessed the grammaticality of each of the sentences that they had just read. To test the influence of secondary tasks, we focused on the obligatory (volitional) and optional (polarity) uses of the subjunctive, both of which have been described in Sections 4.2 and 5.2, respectively.

This chapter is organised as follows. In Section 6.2, we discuss how task design might impact on L2 processing patterns, focusing in particular on the constructs of implicit and explicit knowledge (Section 6.2.1), especially in the context of grammaticality judgment tasks (Section 6.2.2), and on the depth of L2 processing (Section 6.2.3). Following this, we describe the research questions and predictions in Section 6.3 and the methodology in Section 6.4, providing details about the experimental design, including participants, materials and procedure. In Sections 6.5 and 6.6, we present the results and discussion of these results from the obligatory subjunctive study and the polarity subjunctive study, respectively. In Section 6.7, we summarise the findings from the two studies.

6.2 Theoretical Background

As mentioned in the introduction, recent studies have reported that L2 processing patterns may be modulated by different task demands, such as grammaticality judgment tasks versus comprehension questions (Jackson and Bobb, 2009; Leeser, Brandl and Weissglass, 2011), plausibility judgment tasks versus memory tasks (Williams, 2006), the nature and/or number of comprehension questions (Havik *et al.*, 2009), or reading goals (e.g., for comprehension versus for translation) (Lim and Christianson, 2013b, 2015). The question of why differences in task demands lead to differences in processing outcomes is open to debate. The current sub-section seeks to explore the causes of such differences, focusing on both usage-based and formal-based approaches to language acquisition.

6.2.1 Implicit and Explicit Knowledge of Language

Some researchers have argued that L2 learners are more likely to exhibit target-like processing patterns in form-focused tasks, such as translation or judgment tasks, because these tasks are thought to activate explicit knowledge as opposed to implicit knowledge. The constructs of implicit and explicit knowledge have been widely discussed in both cognitive psychology (Reber, 1976; Dienes and Perner, 1999; Cleeremans and Jiménez, 2002) and usage-based approaches to second language research (Ellis, 1994, 2005; Williams, 2009; Rebuschat, 2013). The fundamental difference between implicit and explicit knowledge is said to relate to a person's awareness of what they know (Godfroid *et al.*, 2015). For example, in the case of explicit (conscious metalinguistic) knowledge, a speaker has an awareness of and can consciously demonstrate their knowledge, whereas in the case of implicit (underlying tacit) knowledge, a speaker may not be conscious of what they know and thus apply this knowledge unconsciously. It is important to note that while formal- and usage-based approaches to language acquisition largely disagree on the mental architecture sub-serving these types of knowledge, both paradigms make a principled distinction between metalinguistic knowledge (explicit/learned) and underlying tacit knowledge (implicit/acquired) (Whong, Marsden and Gil, 2013).

Related to this is how these two types of knowledge are accessed. It has been argued that explicit knowledge of language involves declarative representations (Paradis, 2009; Ullman, 2015) that typically only become activated during controlled processing (Ellis, 2009). Implicit knowledge of language, on the other hand, is said to consist of procedural representations (Paradis, 2009; Ullman, 2015) that are activated during automatic processing (Ellis, 2009). The difference between these two types of processing is said to be attentional in nature (Shiffrin and Schneider, 1977; Bialystok, 1994).

Furthermore, research has demonstrated that a number of factors can influence the type of linguistic knowledge that a speaker activates during language comprehension, including L2 experience and task design (Leeser, Brandl and Weissglass, 2011; Godfroid and Winke, 2015). For example, Godfroid and colleagues (2015) argued that learners whose primary linguistic input is through instructed (classroom) exposure activate predominately explicit knowledge during comprehension; this is thought to be the main source of knowledge to which they have access. However, L2 learners who have largely received naturalistic exposure to the target language are more likely to access implicit knowledge during comprehension.

Task design has also been found to modulate the type of knowledge that L2 learners activate during processing. Several studies have shown that learners are more likely to activate implicit knowledge in tasks that either involve a focus on meaning, as opposed to form, and/or impose time constraints, whereas tasks that focus on (linguistic) form with no time constraints are argued to elicit explicit knowledge (Leeser, Brandl and Weissglass, 2011; Godfroid *et al.*, 2015).

For example, Leeser and colleagues (2011) examined whether secondary tasks (comprehension questions vs. grammaticality judgment tasks) influenced intermediate L2 Spanish learners' processing of noun-adjective gender agreement and subject-verb inversion in *wh*-questions. Findings from a self-paced reading task revealed that the type of secondary task used in sentence processing research can indeed modulate L2 learners' sensitivity to morphosyntactic violations, but that the type of structure also plays a role. In particular, this study showed that L2 learners were only sensitive to noun-adjective agreement violations when reading for judgment, as opposed to comprehension. In contrast, L2 learners did not exhibit sensitivity to subject-verb inversion violations, regardless of task type. Based on these findings, Leeser and colleagues argued that when reading for judgment, L2 learners typically activate metalinguistic knowledge that is independent of the type of linguistic knowledge accessed when simply reading for comprehension.

A study by Godfroid and colleagues (2015) examined the effect of time pressure on L2 English learners' eye movements as they judged sentences with and without time constraints. The results showed that timed and untimed GJTs reflect different constructs, namely implicit and explicit knowledge, respectively. More specifically, the authors argued that these findings indicated differences in the levels of automatic and controlled processing required to engage in timed and untimed tasks.

6.2.2 Grammaticality Judgments: Measures of Implicit or Explicit Knowledge?

Although grammaticality judgment tasks have been widely used in L2 research, there remains a significant lack of consensus surrounding the type of knowledge (implicit, explicit or both) that these tasks elicit, particularly among L2 learners and educated native speakers. Many researchers, particularly within the generative paradigm, have argued that if designed appropriately, grammaticality judgments can tap into L2 learners' underlying linguistic representations (Ionin, 2011; Schütze and Sprouse, 2014). However, others have suggested that these tasks may lead L2 learners to activate metalinguistic knowledge instead to complete them (Bialystok, 1979, 1982; Leeser, Brandl and Weissglass, 2011; Godfroid *et al.*, 2015).

For example, Bialystok (1979, 1982) conducted two ground-breaking studies that examined the influence of different task features and instructions on participants' GJT performance in order to investigate learners' use of explicit and implicit knowledge during grammaticality judgment tasks. The 1979 study focused on the effect of time constraints (3-seconds and 15-seconds) using auditory GJTs with L2 French learners. Given that time constraints did not modulate participants' sensitivity to ungrammatical items, Bialystok concluded that globally, auditory GJTs encourage the activation of implicit knowledge. Bialystok did, however, find that tasks requiring a focus on form (e.g., instructing participants to identify the linguistic anomaly in the stimuli) encouraged the use of explicit knowledge. Task modality also played an important role, with written GJTs promoting controlled knowledge retrieval unlike auditory GJTs.

A number of more recent studies have investigated how specific task features, such as item grammaticality and time constraints, can modulate the type of knowledge that L2 learners activate in order to make grammaticality judgments (Leeser, Brandl and Weissglass, 2011;

Godfroid *et al.*, 2015). This research has largely shown that time constraints, as opposed to item grammaticality, play a deterministic role in modulating L2 performance on GJTs. In particular, timed and untimed written GJTs are argued to reflect implicit and explicit knowledge, respectively, a phenomenon discussed in more depth in the preceding section. A notable exception to this finding is a study by Gutiérrez (2013) which found that item grammaticality, and not time constraints, modulated performance on GJTs, and thus suggested that grammatical items represented a measure of implicit knowledge and ungrammatical items a measure of explicit knowledge.

6.2.3 Depth of L2 Morphology Processing

It is important to recognise that not all L2 researchers agree on the logic of the distinction between explicit and implicit knowledge. VanPatten and Rothman (2015, p. 92), for example, argue that the distinction between implicit and explicit knowledge seems illogical "from a generative conceptualisation of what grammar/language is". In particular, they argue that there is "a difference between language learning and linguistic acquisition (in the sense described by Krashen (1981)" but not necessarily an interface between them. As such, true acquisition of a linguistic system must be fundamentally implicit because the process of structure building occurs "as a by-product of first processing the linguistic input itself (e.g., Carroll, 2001; Gregg, 2003; VanPatten, 2011)" (VanPatten and Rothman, 2015, p. 93). The authors therefore encourage a move away from knowledge-testing and an increased focus on the interface between knowledge and processing using online methods, such as eye-tracking, self-paced reading and EEG/ERPs, in order to investigate acquisition-as-processing itself, rather than simply the processing outcomes of acquisition.

It could therefore be argued that differences in processing patterns between tasks do not result from the activation of implicit or explicit knowledge, but rather from the depth of processing. In particular, form-focused tasks, such as grammaticality judgment tasks, are said to direct L2 learners' attention to grammatical form, which in turn leads to longer reading times due to deeper processing. For example, Lim and Christianson (2015) investigated how reading goals (reading for comprehension vs. reading for translation) modulates L2 sensitivity to agreement violations and noun phrase match/mismatch. Findings from an eye-tracking experiment revealed that when required to translate English sentences into Korean, L2 learners demonstrated increased sensitivity to agreement violations, particularly at higher proficiency levels. Lim and Christianson interpreted these findings as indicative of quantitative, rather than qualitative, differences in processing between L1 and L2 speakers, and that more cognitively demanding form-focused tasks may result in deeper, more target-like processing patterns.

Similarly, it has been reported that reading goals also influence the depth of L1 processing. Swets *et al.* (2008) found that the frequency and type of comprehension questions can impact on the depth of processing and final interpretation when resolving syntactic ambiguities. More specifically, this study showed that questions specific to relative clause interpretation, as opposed to superficial questions, led to longer reading times. The authors interpreted these findings as indicating good-enough language processing (Christianson *et al.*, 2001; Ferreira, Bailey and Ferraro, 2002; Ferreira and Patson, 2007), to the extent that readers rely on both syntactic and semantic processing to parse linguistic stimuli, but sometimes not all the available morphosyntactic information is processed.

6.3 Research Questions and Predictions

Based on the literature reviewed above, the current study sought to address whether differences in task demands and thus reading goals influence L2 sensitivity to morphosyntactic violations, and to what extent structure type modulates this pattern. In order to explore this research agenda, we reanalysed data from Study 1 (obligatory subjunctives) and Study 2 (polarity subjunctives). The first reanalysis focused on sensitivity to mood–modality mismatches in obligatory contexts, combining data from Experiments 1a and 1b; the second examined sensitivity to discourse-pragmatic constraints in polarity contexts, merging data from Study 2a and 2b. With this in mind, we advance the following research questions and predictions:

- Does the type of secondary task (reading for comprehension vs. reading for judgment) modulate L2 sensitivity to morphosyntactic violations?
- Does structure type (obligatory subjunctive vs. polarity subjunctive) influence the effect that differences in reading goals and/or task demands has on L2 sensitivity to morphosyntactic violations?
- Does proficiency play a role in this respect?

Based on the aforementioned research questions and the literature discussed in the preceding sections, we anticipate that L2 processing patterns will be more target-like in the judgment task as opposed to the comprehension task. Furthermore, we predict that while this finding will be consistent across structure types, the effect will be more distinct with obligatory than polarity subjunctives. This prediction assumes that sensitivity to the discourse-pragmatic constraints will be limited at best, in light of previous research attesting to the acquisitional difficulty of the subjunctive in polarity contexts. Finally, we expect proficiency to play a significant role, regardless of structure or task type, with more target-like sensitivity surfacing at higher proficiency levels.

6.4 Method

6.4.1 Participants

Participants were identical to those in Study 1 and 2. To recap, we tested a total of 75 participants, including 30 L1 speakers of French and a test group of 45 (English-speaking) L2 learners of French. All participants reported normal or corrected-to-normal vision and were recruited from the student population at the University of Southampton. They were paid £10.00 for their participation in this study. Any participant who reported more than one L1 or whose L1 was not English or French were excluded from the analysis. In light of previous research into the effects of educational attainment and other socioeconomic factors on language outcomes (Pakulak and Neville, 2010; Mulder and Hulstijn, 2011), we restricted recruitment to participants from similar educational and socioeconomic backgrounds to the L2 group (i.e., students or recent graduates with at least two years of university education).

6.4.2 Materials and Design

The materials and the experimental design used in Study 3a (Obligatory) and 3b (Polarity) were identical to those used in Study 1 and 2, respectively. For further details of the materials and the experimental design, see Sections 4.5 and 5.4 of this thesis, respectively.

6.4.3 Procedure

6.4.3.1 Study 3a (Obligatory)

Both tasks were identical with respect to stimuli. The only difference between the two studies was the type of secondary task used. The order in which the two parts of the experiments took place was fixed and thus not counterbalanced, with the judgment task always following the comprehension task.[37] In Task 1, participants were instructed to silently read each sentence for comprehension at their normal reading rate (for no more than 10 seconds) and then answer a

---

[37] We acknowledge that this will have led to repetition effects and quite possibly shorted overall reading times in the judgment task due to repeated exposure to experimental stimuli. However, we consciously adopted this order since the comprehension task was designed to reflect more closely natural reading, allowing participants to read for comprehension rather than for form. In this sense, it was designed to be the more implicit of the two tasks. Had the judgment task preceded the comprehension task, it is likely that the participants would have focused more on grammatical form in the comprehension task, which would have contradicted the overall goal of the comprehension task.

comprehension question (within 15 seconds) on 50% of the trials. Comprehension questions focused on the content of the sentence, not the form, and required a yes-no push-button response. In Task 2, participants were instructed to silently read each sentence for a maximum of 20 seconds and then judge how natural the sentence sounded by selecting a number between 1 and 7, where '1' indicated that the sentence was completely unacceptable and '7' completely acceptable. An acceptability judgment task was included in this study to probe speakers' offline knowledge of the subjunctive.

Both experiments started with a nine-point calibration procedure followed by six practice trials and a total of 36 experimental trials, which were presented in an individually randomized order and intermixed with 36 filler items. At the start of each trial, a fixation point appeared at the location of the first letter of the sentence. Once the participant's fixation on this point was stable, the sentence was presented. Where necessary, recalibration was performed to compensate for any drift in calibration. We aimed for a mean calibration of 0.5 degrees of visual angle. The entire experiment lasted approximately 10 minutes in total.

6.4.3.2 Study 3b (Polarity)

Both experiments were identical with respect to stimuli. Again, the only difference between the two studies was the type of secondary task used. As with Study 3b (Obligatory), the order in which the two parts of the experiments took place was fixed and thus not counterbalanced, with the judgment task always following the comprehension task. In Task 1, participants were instructed to look at the contextualising picture (for a maximum of 15 seconds), silently read each sentence for comprehension at their normal reading rate (for no more than 10 seconds) and then answer a comprehension question (within 15 seconds) on 50% of all trials. Comprehension questions focused on whether the contextualising picture matched the sentence and required a yes-no push-button response. In Task 2, participants were asked to look at the contextualising picture (for a maximum of 15 seconds), silently read each sentence for a maximum of 20 seconds and then judge how natural the sentence sounded by selecting a number between 1 and 7, where '1' indicated that the sentence was completely unacceptable and '7' completely acceptable. An acceptability judgment task was included in this study to probe speakers' offline knowledge of the subjunctive.

Both experiments started with a nine-point calibration procedure followed by six practice trials and a total of 32 experimental trials, which were presented in an individually randomized order and intermixed with 32 filler items. At the start of each trial, a fixation point appeared at the location of the first letter of the sentence. Once the participant's fixation on this point was stable, the picture was presented. Where necessary, recalibration was performed to compensate for any drift in calibration. We aimed for a mean calibration of 0.5 degrees of visual angle. The entire experiment lasted approximately 10 minutes in total.

Appendix A

Table 6.1 Mixed-effects models for L1 global (obligatory subjunctive)

		b	*SE*	*z*
Mean Fixation Count	(Intercept)	14.80	0.54	27.53
	Mood: Subjunctive vs. indicative	-0.06	0.09	-0.62
	Semantic: Desire vs. Directive	-0.44	0.26	-1.67
	Semantic: Emotive-factive vs. Directive	-0.02	0.26	-0.08
	Task (Judgment vs. Comprehension)	1.56	0.09	17.02
	Sentence Length (Centred)	1.83	0.19	9.66
	Trial Number	-0.04	0.00	-7.91
	Mood × Semantic (Desire vs. Directive)	-0.10	0.13	-0.80
	Mood × Semantic (Emotive-Factive vs. Directive)	-0.03	0.13	-0.20
	Mood × Task	0.42	0.09	4.60
	Semantic (Desire vs. Directive) × Task	0.14	0.13	1.05
	Semantic (Emotive-Factive vs. Directive) × Task	0.14	0.13	1.11
	Mood × Semantic (Desire vs. Directive) × Task	0.26	0.13	2.03
	Mood × Semantic (Emotive-Factive vs. Directive) × Task	-0.13	0.13	-1.04
Mean Fixation Duration	(Intercept)	5.47	0.02	329.61
	Mood: Subjunctive vs. indicative	-0.02	0.00	-6.75
	Semantic: Desire vs. Directive	-0.01	0.00	-1.25
	Semantic: Emotive-factive vs. Directive	0.00	0.00	-0.20
	Task (Judgment vs. Comprehension)	0.00	0.00	0.60
	Sentence Length (Centred)	-0.01	0.00	-2.49
	Trial Number	0.00	0.00	-3.29
	Mood × Semantic (Desire vs. Directive)	0.00	0.00	0.29
	Mood × Semantic (Emotive-Factive vs. Directive)	0.00	0.00	0.27
	Mood × Task	-0.01	0.00	-2.99
	Semantic (Desire vs. Directive) × Task	0.01	0.00	1.53
	Semantic (Emotive-Factive vs. Directive) × Task	0.00	0.00	-0.36
	Mood × Semantic (Desire vs. Directive) × Task	0.00	0.00	0.29
	Mood × Semantic (Emotive-Factive vs. Directive) × Task	-0.01	0.00	-1.98
Total Sentence Reading Time	(Intercept)	8.44	0.04	213.16
	Mood: Subjunctive vs. indicative	-0.03	0.01	-4.82
	Semantic: Desire vs. Directive	-0.03	0.02	-1.87
	Semantic: Emotive-factive vs. Directive	0.00	0.02	0.27
	Task (Judgment vs. Comprehension)	0.08	0.01	12.00
	Sentence Length (Centred)	0.12	0.01	10.17
	Trial Number	0.00	0.00	-12.83
	Mood × Semantic (Desire vs. Directive)	0.00	0.01	0.14
	Mood × Semantic (Emotive-Factive vs. Directive)	0.00	0.01	-0.36
	Mood × Task	0.02	0.01	2.55
	Semantic (Desire vs. Directive) × Task	0.02	0.01	2.36
	Semantic (Emotive-Factive vs. Directive) × Task	0.01	0.01	0.97
	Mood × Semantic (Desire vs. Directive) × Task	0.02	0.01	2.06
	Mood × Semantic (Emotive-Factive vs. Directive) × Task	-0.02	0.01	-2.41

Note: Shaded cells denote statistical significance (p < .05).

Table 6.2 Mixed-effects models for L2 global (obligatory subjunctive)

	Mean Fixation Count			Mean Fixation Duration			Total Sentence Reading Time		
	b	*SE*	*z*	*b*	*SE*	*z*	*b*	*SE*	*z*
(Intercept)	18.27	1.41	13.00	5.51	0.02	265.57	8.65	0.07	118.13
Mood: Subjunctive vs. Indicative	0.16	0.19	0.84	0.00	0.00	-1.47	-0.01	0.01	-0.51
Semantic: Desire vs. Directive	-0.84	0.40	-2.13	0.00	0.01	-0.46	-0.06	0.02	-2.38
Semantic: Emotive-Factive vs. Directive	-0.37	0.39	-0.95	-0.01	0.01	-0.93	-0.01	0.02	-0.43
Task (Judgment vs. Comprehension)	1.63	0.19	8.78	0.00	0.00	0.44	0.09	0.01	9.25
Proficiency (Centred)	-2.04	1.56	-1.30	-0.02	0.02	-0.96	-0.12	0.08	-1.43
Sentence Length (Centred)	1.97	0.24	8.28	0.00	0.00	1.19	0.10	0.02	6.84
Trial Number	-0.03	0.01	-6.62	0.00	0.00	-1.01	0.00	0.00	-8.34
Mood × Semantic (Desire vs. Directive)	0.13	0.26	0.49	0.01	0.00	1.35	0.00	0.01	0.02
Mood × Semantic (Emotive-Factive vs. Directive)	-0.02	0.26	-0.07	0.00	0.00	-1.00	0.01	0.01	0.47
Mood × Task	0.19	0.19	1.02	-0.01	0.00	-2.18	0.00	0.01	-0.03
Semantic (Desire vs. Directive) × Task	0.08	0.26	0.32	0.00	0.00	0.41	0.02	0.01	1.08
Semantic (Emotive-Factive vs. Directive) × Task	0.30	0.26	1.15	-0.01	0.00	-1.36	0.02	0.01	1.67
Mood × Prof.	-0.06	0.21	-0.27	0.00	0.00	-0.80	-0.01	0.01	-1.18
Semantic (Desire vs. Directive) × Prof.	-0.49	0.30	-1.65	0.00	0.01	-0.22	-0.03	0.02	-2.12
Semantic (Emotive-Factive vs. Directive) × Prof.	-0.24	0.30	-0.80	0.00	0.01	-0.18	0.00	0.02	0.27
Task × Prof.	-0.25	0.21	-1.15	0.00	0.00	-0.66	0.02	0.01	1.51
Mood × Semantic (Desire vs. Directive) × Task	0.29	0.26	1.12	0.00	0.00	-0.15	0.01	0.01	0.51
Mood × Semantic (Emotive-Factive vs. Directive) × Task	-0.37	0.26	-1.39	-0.01	0.00	-1.26	-0.02	0.01	-1.24
Mood × Semantic (Desire vs. Directive) × Prof.	0.45	0.30	1.51	0.01	0.01	1.29	0.02	0.02	1.04
Mood × Semantic (Emotive-Factive vs. Directive) × Prof.	-0.21	0.30	-0.69	-0.01	0.01	-1.20	0.00	0.02	-0.27
Mood × Task × Prof.	0.16	0.21	0.74	0.00	0.00	-0.84	0.00	0.01	0.38
Semantic (Desire vs. Directive) × Task × Prof.	-0.06	0.30	-0.19	0.00	0.01	-0.49	0.00	0.02	-0.06
Semantic (Emotive-Factive vs. Directive) × Task × Prof.	0.29	0.30	0.96	-0.01	0.01	-1.20	0.02	0.02	1.28
Mood × Semantic (Desire vs. Directive) × Task × Prof.	0.39	0.30	1.30	0.01	0.01	1.32	0.02	0.02	1.00
Mood × Semantic (Emotive-Factive vs. Directive) × Task × Prof.	-0.18	0.30	-0.61	-0.01	0.01	-2.61	-0.01	0.02	-0.57

Note: Shaded cells denote statistical significance ($p < .05$).

6.5 Results: Study 3a (Obligatory)

6.5.1 Data Preparation and Analyses

Data preparation and analyses for Study 3a were almost identical to those of Study 1a and 1b. The only difference was the inclusion of task as a categorical fixed effect and the exclusion of residence abroad. Sum contrasts were coded for task, exploring the difference between judgment comprehension (judgment -1, comprehension 1).

6.5.2 Global Measures

6.5.2.1 L1 Group

To examine the overall effect of mood and its interaction with matrix semantic property and task on reading behaviour, we calculated the following global measures: mean fixation duration, mean fixation count and total sentence reading time. The fixed effect estimates for global measures are shown in Table 6.1.

Although we did not find a significant main effect of mood for mean fixation count, there was a significant two-way interaction between mood and task, such that in the comprehension task, indicative sentences received more fixations on average than subjunctive sentences, whereas the reverse was true in the judgment task. That is, subjunctive sentences received more fixations on average than indicative sentences.

For mean fixation duration, there was a significant main effect of mood. Here, mean fixation duration was longer for indicative than subjunctive sentences. Several significant interactions (i.e., Mood × Task and Mood × Semantic (Emotive-Factive vs. Directive) × Task) suggest this pattern was modulated by task and verb type. Although mean fixation duration was higher for indicative than subjunctive sentences across tasks, this effect was magnified in the judgment task, particularly with emotive-factive verbs (compared with directive verbs).

For total sentence reading time, there was a significant main effect of mood, such that total sentence reading time was longer for indicative than subjunctive sentences. Several significant interactions (i.e., Mood × Task, Mood × Semantic (Emotive-Factive vs. Directive) × Task and Mood × Semantic (Desire vs. Directive) × Task) indicate that this pattern was modulated by task and matrix semantic property. More specifically, the difference between indicative and subjunctive sentences was particularly strong in the comprehension task. Although total sentence reading time was consistently longer for indicative (than subjunctive) sentences, this was not always the case in the judgment task. In particular, we found that total sentence reading time was marginally higher for subjunctive than indicative sentences with desire and emotive-factive verbs.

6.5.2.2 L2 Group

To examine L2 learners' sensitivity to mood and the influence of matrix semantic property, task, and proficiency on reading behaviour, we calculated the following global measures: mean fixation duration, mean fixation count and total sentence reading time. The fixed effect estimates for the global measures are shown in Table 6.2.

For mean fixation count and total sentence reading time, we did not find a significant main effect of mood or task. For mean fixation duration, however, the following interactions were significant: Mood × Task and Mood × Semantic (Emotive-Factive vs. Directive) × Task × Proficiency. The two-way interaction between mood and task revealed that mean fixation duration was longer for indicative (than subjunctive) sentences in the judgment task, but longer for subjunctive (than indicative) sentences in the comprehension task.

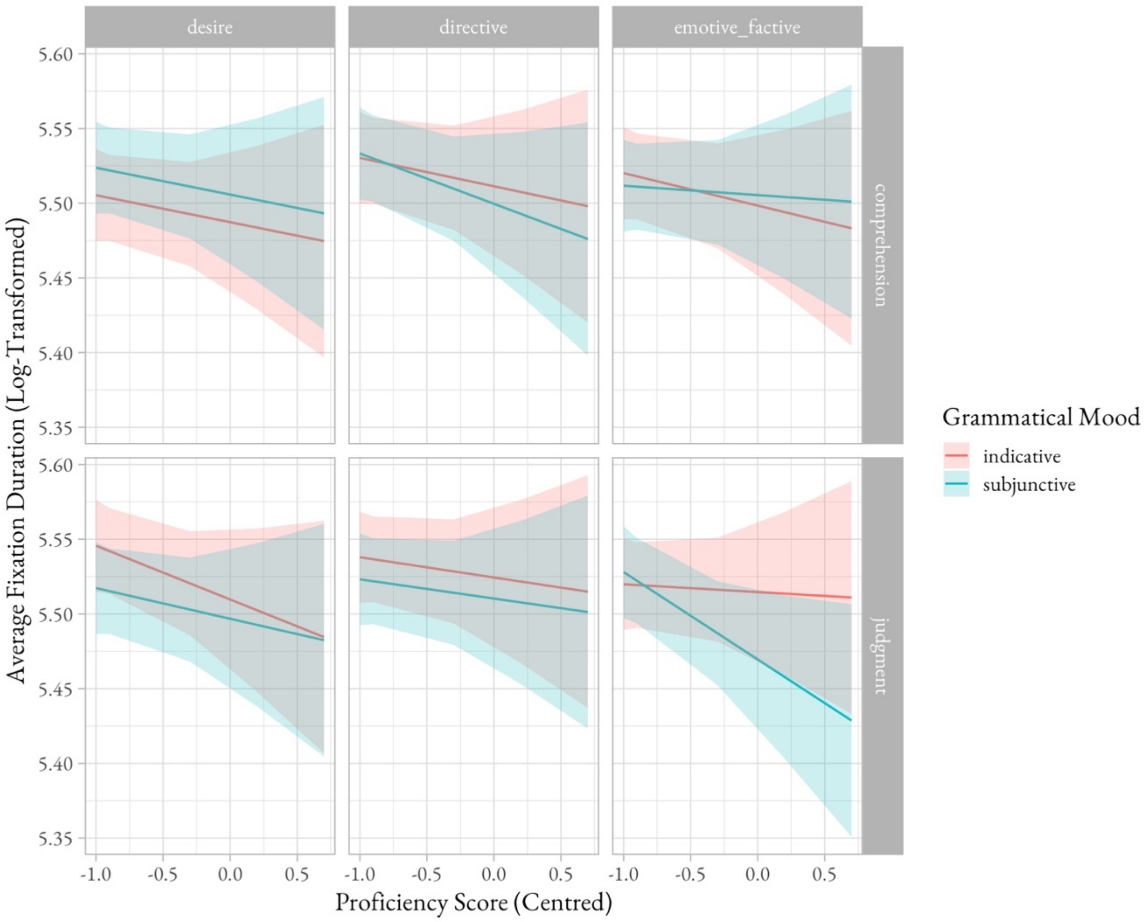


Figure 6.1 Predicted total sentence reading time

Figure 6.1 visualises the four-way interaction between mood, matrix semantic property (emotive-factive vs. directive), task and proficiency. It shows that mean fixation duration in the comprehension task was numerically longer for indicative than subjunctive sentences with directive verbs at higher proficiency, but numerically longer for subjunctive than indicative sentences with both desire[38] and emotive-factive verbs.[39] In contrast, mean fixation duration in the judgment task was, for the most part, numerically longer for indicative than subjunctive sentences across matrix semantic properties. Note, however, that the difference was minimal with desire verbs at the highest proficiency levels but strengthened as a function of proficiency with emotive-factive verbs to the extent that the difference in mean fixation duration between indicative and subjunctive sentences reached significance.

[38] It is interesting that we did not find a significant difference between desire and directive verbs in this context.

[39] We acknowledge that at the lower proficiency levels, mean fixation duration was in fact longer for indicative than subjunctive sentences, but this trend quickly reversed as proficiency increased.

Table 6.3 Mixed-effects models for L1 local (obligatory subjunctive)

		Critical Region			Spillover Region		
		b	*SE*	*t/z*	*b*	*SE*	*t/z*
First Fixation Duration	(Intercept)	5.46	0.03	186.83	5.43	0.02	229.12
	Mood: Subjunctive vs. Indicative	-0.10	0.03	-3.58	-0.01	0.01	-0.94
	Semantic: Desire vs. Directive	-0.01	0.01	-0.91	0.02	0.01	1.28
	Semantic: Emotive-Factive vs. Directive	0.01	0.01	0.60	0.00	0.01	-0.31
	Task (Judgment vs. Comprehension)	0.01	0.01	1.43	0.00	0.01	0.67
	Sentence Length (Centred)	-0.06	0.03	-2.31	0.01	0.01	0.58
	Trial Number	0.00	0.00	-0.51	0.00	0.00	1.04
	Mood × Semantic (Desire vs. Directive)	-0.03	0.01	-1.98	0.00	0.01	0.06
	Mood × Semantic (Emotive-Factive vs. Directive)	0.01	0.01	1.01	0.01	0.01	0.77
	Mood × Task	-0.01	0.01	-1.27	0.00	0.01	0.05
	Semantic (Desire vs. Directive) × Task	0.00	0.01	0.29	0.01	0.01	1.31
	Semantic (Emotive-Factive vs. Directive) × Task	-0.01	0.01	-1.16	-0.01	0.01	-1.10
	Mood × Semantic (Desire vs. Directive) × Task	0.00	0.01	0.03	-0.01	0.01	-1.38
	Mood × Semantic (Emotive-Factive vs. Directive) × Task	0.01	0.01	0.54	-0.01	0.01	-0.61
Gaze Duration	(Intercept)	5.49	0.03	175.62	5.82	0.04	141.43
	Mood: Subjunctive vs. Indicative	-0.10	0.03	-3.47	-0.02	0.01	-1.77
	Semantic: Desire vs. Directive	-0.02	0.02	-1.51	0.07	0.02	3.11
	Semantic: Emotive-Factive vs. Directive	0.01	0.02	0.83	-0.06	0.02	-2.76
	Task (Judgment vs. Comprehension)	0.02	0.01	1.66	0.02	0.01	1.87
	Sentence Length (Centred)	-0.06	0.03	-2.15	0.17	0.02	10.62
	Trial Number	0.00	0.00	-0.40	0.00	0.00	-1.03
	Mood × Semantic (Desire vs. Directive)	-0.03	0.01	-1.79	0.01	0.02	0.77
	Mood × Semantic (Emotive-Factive vs. Directive)	0.01	0.01	0.60	0.00	0.02	-0.14
	Mood × Task	-0.01	0.01	-1.17	0.00	0.01	0.32
	Semantic (Desire vs. Directive) × Task	0.00	0.01	0.14	0.00	0.02	-0.26
	Semantic (Emotive-Factive vs. Directive) × Task	0.00	0.01	-0.32	0.01	0.02	0.43
	Mood × Semantic (Desire vs. Directive) × Task	0.01	0.01	0.48	-0.01	0.02	-0.45
	Mood × Semantic (Emotive-Factive vs. Directive) × Task	0.00	0.01	-0.26	-0.03	0.02	-1.85
Go-Past Time	(Intercept)	5.53	0.03	168.92	6.06	0.04	135.45
	Mood: Subjunctive vs. Indicative	-0.10	0.03	-2.96	-0.08	0.01	-7.16
	Semantic: Desire vs. Directive	-0.02	0.02	-1.17	0.06	0.04	1.65
	Semantic: Emotive-Factive vs. Directive	0.01	0.02	0.72	-0.07	0.04	-1.90
	Task (Judgment vs. Comprehension)	0.03	0.01	3.08	0.07	0.01	6.44
	Sentence Length (Centred)	-0.06	0.03	-1.62	0.23	0.03	8.49
	Trial Number	0.00	0.00	-0.55	0.00	0.00	-3.69
	Mood × Semantic (Desire vs. Directive)	-0.02	0.02	-1.22	0.01	0.02	0.70
	Mood × Semantic (Emotive-Factive vs. Directive)	0.01	0.02	0.61	0.01	0.02	0.34
	Mood × Task	-0.02	0.01	-1.81	0.00	0.01	0.12
	Semantic (Desire vs. Directive) × Task	0.01	0.02	0.68	-0.03	0.02	-1.64
	Semantic (Emotive-Factive vs. Directive) × Task	-0.01	0.01	-0.68	0.02	0.02	1.32
	Mood × Semantic (Desire vs. Directive) × Task	0.01	0.02	0.71	0.00	0.02	0.22
	Mood × Semantic (Emotive-Factive vs. Directive) × Task	-0.01	0.01	-0.52	-0.04	0.02	-2.32
Total Reading Time	(Intercept)	5.75	0.05	126.05	6.39	0.05	139.26
	Mood: Subjunctive vs. Indicative	-0.13	0.05	-2.55	-0.09	0.01	-7.34
	Semantic: Desire vs. Directive	-0.04	0.03	-1.41	0.06	0.02	3.36
	Semantic: Emotive-Factive vs. Directive	0.01	0.03	0.39	-0.07	0.02	-4.00
	Task (Judgment vs. Comprehension)	0.06	0.01	4.33	0.10	0.01	8.72
	Sentence Length (Centred)	-0.06	0.05	-1.24	0.24	0.01	20.49
	Trial Number	0.00	0.00	-2.21	0.00	0.00	-6.27
	Mood × Semantic (Desire vs. Directive)	-0.02	0.02	-0.82	0.01	0.02	0.75
	Mood × Semantic (Emotive-Factive vs. Directive)	0.02	0.02	1.16	-0.01	0.02	-0.68
	Mood × Task	-0.02	0.01	-1.13	0.02	0.01	1.93
	Semantic (Desire vs. Directive) × Task	-0.01	0.02	-0.67	-0.01	0.02	-0.33
	Semantic (Emotive-Factive vs. Directive) × Task	0.02	0.02	0.79	0.03	0.02	1.72
	Mood × Semantic (Desire vs. Directive) × Task	0.05	0.02	2.29	0.02	0.02	1.07
	Mood × Semantic (Emotive-Factive vs. Directive) × Task	-0.02	0.02	-0.82	-0.04	0.02	-2.33
Regression-In Probability	(Intercept)	-0.44	0.14	-3.18	-0.80	0.12	-6.89
	Mood: Subjunctive vs. Indicative	-0.69	0.17	-4.13	-0.10	0.05	-1.84

	Semantic: Desire vs. Directive	0.19	0.09	2.02	0.00	0.08	0.04
	Semantic: Emotive-Factive vs. Directive	-0.10	0.09	-1.12	-0.14	0.08	-1.72
	Task (Judgment vs. Comprehension)	0.31	0.06	4.98	0.12	0.05	2.15
	Sentence Length (Centred)	0.01	0.17	0.06	0.05	0.05	0.88
	Trial Number	0.00	0.00	-0.39	-0.01	0.00	-4.35
	Mood × Semantic (Desire vs. Directive)	-0.04	0.09	-0.42	-0.04	0.08	-0.53
	Mood × Semantic (Emotive-Factive vs. Directive)	0.00	0.09	-0.05	-0.10	0.08	-1.26
	Mood × Task	0.13	0.06	2.07	0.06	0.05	1.09
	Semantic (Desire vs. Directive) × Task	-0.04	0.09	-0.40	0.08	0.08	1.08
	Semantic (Emotive-Factive vs. Directive) × Task	-0.05	0.09	-0.51	0.01	0.08	0.09
	Mood × Semantic (Desire vs. Directive) × Task	0.18	0.09	1.95	0.04	0.08	0.51
	Mood × Semantic (Emotive-Factive vs. Directive) × Task	-0.10	0.09	-1.19	-0.03	0.08	-0.41
Regression-Out Probability	(Intercept)	-2.55	0.24	-10.65	-0.82	0.11	-7.31
	Mood: Subjunctive vs. Indicative	-0.08	0.29	-0.29	-0.42	0.05	-8.34
	Semantic: Desire vs. Directive	0.12	0.17	0.70	0.11	0.07	1.55
	Semantic: Emotive-Factive vs. Directive	0.03	0.16	0.22	-0.04	0.07	-0.56
	Task (Judgment vs. Comprehension)	0.33	0.11	2.86	0.25	0.05	4.87
	Sentence Length (Centred)	-0.05	0.29	-0.18	0.21	0.05	4.22
	Trial Number	0.00	0.00	0.00	0.00	0.00	-0.46
	Mood × Semantic (Desire vs. Directive)	0.19	0.17	1.15	-0.11	0.07	-1.55
	Mood × Semantic (Emotive-Factive vs. Directive)	0.13	0.16	0.83	0.07	0.07	0.95
	Mood × Task	-0.22	0.11	-1.93	-0.01	0.05	-0.27
	Semantic (Desire vs. Directive) × Task	0.23	0.17	1.40	-0.13	0.07	-1.86
	Semantic (Emotive-Factive vs. Directive) × Task	-0.23	0.16	-1.45	0.04	0.07	0.55
	Mood × Semantic (Desire vs. Directive) × Task	-0.12	0.17	-0.75	-0.03	0.07	-0.45
	Mood × Semantic (Emotive-Factive vs. Directive) × Task	-0.05	0.16	-0.31	0.06	0.07	0.81

Note: Shaded cells denote statistical significance (p < .05).

Table 6.4 Mixed-effects models for L2 local (obligatory subjunctive)

		Critical region			Spillover region		
		b	*SE*	*t*/*z*	*b*	*SE*	*t*/*z*
First Fixation Duration	(Intercept)	5.47	0.03	182.48	5.53	0.03	196.82
	Mood: Subjunctive vs. Indicative	-0.07	0.03	-2.57	-0.01	0.01	-1.16
	Semantic: Desire vs. Directive	0.02	0.02	1.04	0.02	0.02	0.82
	Semantic: Emotive-Factive vs. Directive	-0.01	0.02	-0.52	-0.03	0.02	-1.47
	Task (Judgment vs. Comprehension)	0.00	0.01	-0.19	0.00	0.01	-0.17
	Proficiency (Centred)	-0.02	0.03	-0.51	-0.01	0.03	-0.48
	Sentence Length (Centred)	-0.07	0.02	-2.77	0.01	0.01	0.98
	Trial Number	0.00	0.00	1.86	0.00	0.00	0.44
	Mood × Semantic (Desire vs. Directive)	0.01	0.02	0.55	0.01	0.01	0.54
	Mood × Semantic (Emotive-Factive vs. Directive)	-0.02	0.02	-0.96	0.00	0.02	-0.31
	Mood × Task	-0.03	0.01	-2.42	-0.01	0.01	-1.12
	Semantic (Desire vs. Directive) × Task	0.00	0.02	-0.06	0.02	0.01	1.12
	Semantic (Emotive-Factive vs. Directive) × Task	0.01	0.02	0.49	-0.03	0.02	-1.84
	Mood × Prof.	-0.02	0.01	-1.61	-0.02	0.01	-1.79
	Semantic (Desire vs. Directive) × Prof.	0.01	0.02	0.61	0.01	0.02	0.62
	Semantic (Emotive-Factive vs. Directive) × Prof.	0.00	0.02	-0.20	-0.01	0.02	-0.57
	Task × Prof.	0.00	0.01	-0.03	-0.02	0.01	-1.36
	Mood × Semantic (Desire vs. Directive) × Task	-0.01	0.02	-0.81	0.00	0.01	-0.18
	Mood × Semantic (Emotive-Factive vs. Directive) × Task	0.00	0.02	-0.24	0.00	0.02	0.26
	Mood × Semantic (Desire vs. Directive) × Prof.	0.01	0.02	0.37	0.03	0.02	1.58
	Mood × Semantic (Emotive-Factive vs. Directive) × Prof.	-0.01	0.02	-0.51	-0.01	0.02	-0.40
	Mood × Task × Prof.	-0.01	0.01	-0.88	-0.01	0.01	-1.22
	Semantic (Desire vs. Directive) × Task × Prof.	0.01	0.02	0.26	0.02	0.02	0.89
	Semantic (Emotive-Factive vs. Directive) × Task × Prof.	0.00	0.02	-0.22	-0.03	0.02	-1.65
	Mood × Semantic (Desire vs. Directive) × Task × Prof.	-0.03	0.02	-1.55	0.00	0.02	0.25
	Mood × Semantic (Emotive-Factive vs. Directive) × Task × Prof.	0.00	0.02	0.18	-0.01	0.02	-0.46
Gaze Duration	(Intercept)	5.50	0.03	160.70	5.95	0.05	122.88
	Mood: Subjunctive vs. Indicative	-0.06	0.03	-2.10	0.01	0.02	0.74
	Semantic: Desire vs. Directive	0.02	0.02	0.95	0.06	0.04	1.41
	Semantic: Emotive-Factive vs. Directive	-0.02	0.02	-0.74	-0.11	0.04	-2.73
	Task (Judgment vs. Comprehension)	0.01	0.01	0.51	0.00	0.02	0.10
	Proficiency (Centred)	-0.03	0.03	-0.84	-0.07	0.04	-1.50
	Sentence Length (Centred)	-0.07	0.03	-2.35	0.18	0.02	7.52
	Trial Number	0.00	0.00	1.33	0.00	0.00	-1.26
	Mood × Semantic (Desire vs. Directive)	0.00	0.02	-0.13	0.02	0.02	0.63
	Mood × Semantic (Emotive-Factive vs. Directive)	-0.02	0.02	-0.90	0.02	0.02	0.91
	Mood × Task	-0.04	0.01	-2.61	-0.03	0.02	-1.95
	Semantic (Desire vs. Directive) × Task	0.00	0.02	0.19	-0.02	0.02	-0.85
	Semantic (Emotive-Factive vs. Directive) × Task	0.01	0.02	0.29	0.00	0.02	0.02
	Mood × Prof.	-0.03	0.02	-1.86	-0.02	0.02	-0.86
	Semantic (Desire vs. Directive) × Prof.	0.01	0.02	0.60	0.01	0.03	0.32
	Semantic (Emotive-Factive vs. Directive) × Prof.	-0.01	0.02	-0.45	-0.04	0.03	-1.62
	Task × Prof.	0.00	0.02	0.26	-0.02	0.02	-0.81
	Mood × Semantic (Desire vs. Directive) × Task	-0.03	0.02	-1.28	0.00	0.02	0.03
	Mood × Semantic (Emotive-Factive vs. Directive) × Task	0.01	0.02	0.70	0.00	0.02	-0.03
	Mood × Semantic (Desire vs. Directive) × Prof.	0.00	0.02	0.16	0.01	0.03	0.29
	Mood × Semantic (Emotive-Factive vs. Directive) × Prof.	-0.01	0.02	-0.52	0.03	0.03	1.23
	Mood × Task × Prof.	-0.02	0.02	-1.06	-0.04	0.02	-2.19
	Semantic (Desire vs. Directive) × Task × Prof.	0.00	0.02	-0.03	0.00	0.03	0.18
	Semantic (Emotive-Factive vs. Directive) × Task × Prof.	0.00	0.02	0.09	-0.01	0.03	-0.51
	Mood × Semantic (Desire vs. Directive) × Task × Prof.	-0.04	0.02	-1.77	0.01	0.03	0.24
	Mood × Semantic (Emotive-Factive vs. Directive) × Task × Prof.	0.01	0.02	0.65	0.00	0.03	-0.14
Go-Past Time	(Intercept)	5.57	0.04	141.07	6.19	0.06	106.32
	Mood: Subjunctive vs. Indicative	-0.08	0.03	-2.56	0.01	0.02	0.66
	Semantic: Desire vs. Directive	0.00	0.02	-0.06	0.09	0.05	1.85
	Semantic: Emotive-Factive vs. Directive	0.01	0.02	0.61	-0.09	0.05	-1.85
	Task (Judgment vs. Comprehension)	0.03	0.02	1.75	0.06	0.02	3.89
	Proficiency (Centred)	-0.02	0.04	-0.52	-0.10	0.05	-1.83

		Critical region			Spillover region		
		b	*SE*	*t/z*	*b*	*SE*	*t/z*
	Sentence Length (Centred)	-0.08	0.03	-2.78	0.24	0.03	7.41
	Trial Number	0.00	0.00	0.06	0.00	0.00	-2.53
	Mood × Semantic (Desire vs. Directive)	-0.02	0.02	-0.82	-0.01	0.02	-0.64
	Mood × Semantic (Emotive-Factive vs. Directive)	-0.01	0.02	-0.29	0.02	0.02	0.99
	Mood × Task	-0.02	0.02	-1.02	-0.02	0.02	-1.21
	Semantic (Desire vs. Directive) × Task	-0.01	0.02	-0.46	0.00	0.02	0.01
	Semantic (Emotive-Factive vs. Directive) × Task	0.03	0.02	1.59	0.01	0.02	0.45
	Mood × Prof.	-0.03	0.02	-1.96	-0.01	0.02	-0.29
	Semantic (Desire vs. Directive) × Prof.	0.00	0.03	0.03	0.03	0.03	1.20
	Semantic (Emotive-Factive vs. Directive) × Prof.	0.00	0.02	0.10	-0.03	0.03	-1.16
	Task × Prof.	0.00	0.02	0.08	0.00	0.02	0.07
	Mood × Semantic (Desire vs. Directive) × Task	-0.04	0.02	-1.89	-0.01	0.02	-0.43
	Mood × Semantic (Emotive-Factive vs. Directive) × Task	0.02	0.02	1.05	0.00	0.02	-0.12
	Mood × Semantic (Desire vs. Directive) × Prof.	-0.01	0.03	-0.26	0.00	0.03	0.00
	Mood × Semantic (Emotive-Factive vs. Directive) × Prof.	0.00	0.02	0.01	0.00	0.03	-0.07
	Mood × Task × Prof.	0.01	0.02	0.65	-0.02	0.02	-1.29
	Semantic (Desire vs. Directive) × Task × Prof.	-0.01	0.03	-0.54	0.03	0.03	1.23
	Semantic (Emotive-Factive vs. Directive) × Task × Prof.	0.02	0.02	1.03	0.00	0.03	0.13
	Mood × Semantic (Desire vs. Directive) × Task × Prof.	-0.06	0.03	-2.25	0.00	0.03	-0.07
	Mood × Semantic (Emotive-Factive vs. Directive) × Task × Prof.	0.03	0.02	1.19	0.00	0.03	-0.01
Total Reading Time	(Intercept)	5.82	0.07	86.36	6.62	0.08	87.94
	Mood: Subjunctive vs. Indicative	-0.05	0.06	-0.91	-0.01	0.02	-0.44
	Semantic: Desire vs. Directive	-0.01	0.04	-0.39	0.04	0.03	1.52
	Semantic: Emotive-Factive vs. Directive	0.02	0.04	0.49	-0.13	0.03	-4.98
	Task (Judgment vs. Comprehension)	0.06	0.02	2.83	0.07	0.02	3.64
	Proficiency (Centred)	-0.04	0.07	-0.55	-0.12	0.08	-1.49
	Sentence Length (Centred)	-0.08	0.06	-1.35	0.23	0.01	22.32
	Trial Number	0.00	0.00	0.42	0.00	0.00	-4.79
	Mood × Semantic (Desire vs. Directive)	-0.02	0.03	-0.73	0.03	0.03	1.31
	Mood × Semantic (Emotive-Factive vs. Directive)	-0.02	0.03	-0.58	0.01	0.03	0.45
	Mood × Task	-0.02	0.02	-0.93	0.00	0.02	-0.15
	Semantic (Desire vs. Directive) × Task	0.01	0.03	0.18	0.02	0.03	0.65
	Semantic (Emotive-Factive vs. Directive) × Task	0.03	0.03	1.22	0.01	0.03	0.40
	Mood × Prof.	-0.05	0.02	-2.16	-0.01	0.02	-0.27
	Semantic (Desire vs. Directive) × Prof.	0.00	0.03	-0.01	-0.03	0.03	-1.07
	Semantic (Emotive-Factive vs. Directive) × Prof.	-0.01	0.03	-0.43	-0.04	0.03	-1.18
	Task × Prof.	0.02	0.02	0.67	-0.01	0.02	-0.49
	Mood × Semantic (Desire vs. Directive) × Task	0.02	0.03	0.71	0.03	0.03	1.16
	Mood × Semantic (Emotive-Factive vs. Directive) × Task	-0.01	0.03	-0.40	-0.05	0.03	-1.82
	Mood × Semantic (Desire vs. Directive) × Prof.	-0.01	0.03	-0.29	0.07	0.03	2.24
	Mood × Semantic (Emotive-Factive vs. Directive) × Prof.	-0.01	0.03	-0.22	-0.03	0.03	-0.99
	Mood × Task × Prof.	0.01	0.02	0.57	0.00	0.02	0.11
	Semantic (Desire vs. Directive) × Task × Prof.	0.00	0.03	0.03	0.00	0.03	0.16
	Semantic (Emotive-Factive vs. Directive) × Task × Prof.	0.01	0.03	0.22	0.02	0.03	0.64
	Mood × Semantic (Desire vs. Directive) × Task × Prof.	0.03	0.03	1.03	0.04	0.03	1.31
	Mood × Semantic (Emotive-Factive vs. Directive) × Task × Prof.	-0.02	0.03	-0.58	-0.03	0.03	-1.14
Regression-In Probability	(Intercept)	-0.60	0.12	-4.91	-0.53	0.10	-5.25
	Mood: Subjunctive vs. Indicative	-0.08	0.14	-0.54	-0.02	0.07	-0.24
	Semantic: Desire vs. Directive	0.00	0.12	0.00	-0.04	0.10	-0.44
	Semantic: Emotive-Factive vs. Directive	-0.09	0.11	-0.76	-0.19	0.10	-1.95
	Task (Judgment vs. Comprehension)	0.02	0.08	0.31	0.14	0.07	2.04
	Proficiency (Centred)	-0.15	0.09	-1.63	-0.21	0.08	-2.66
	Sentence Length (Centred)	0.25	0.13	1.99	0.02	0.04	0.56
	Trial Number	0.00	0.00	-0.10	0.00	0.00	-2.39
	Mood × Semantic (Desire vs. Directive)	0.02	0.12	0.15	0.18	0.10	1.79
	Mood × Semantic (Emotive-Factive vs. Directive)	0.03	0.11	0.24	-0.12	0.10	-1.23
	Mood × Task	0.00	0.08	0.04	0.14	0.07	2.07
	Semantic (Desire vs. Directive) × Task	0.11	0.12	0.98	-0.10	0.10	-1.05
	Semantic (Emotive-Factive vs. Directive) × Task	-0.10	0.11	-0.86	0.09	0.10	0.93
	Mood × Prof.	-0.07	0.09	-0.77	0.05	0.08	0.59

		Critical region			Spillover region		
		b	*SE*	*t/z*	*b*	*SE*	*t/z*
	Semantic (Desire vs. Directive) × Prof.	-0.05	0.13	-0.37	-0.21	0.11	-1.88
	Semantic (Emotive-Factive vs. Directive) × Prof.	0.00	0.13	0.02	0.11	0.11	0.96
	Task × Prof.	-0.12	0.09	-1.31	0.03	0.08	0.40
	Mood × Semantic (Desire vs. Directive) × Task	0.20	0.12	1.72	0.07	0.10	0.67
	Mood × Semantic (Emotive-Factive vs. Directive) × Task	-0.29	0.11	-2.58	-0.02	0.10	-0.23
	Mood × Semantic (Desire vs. Directive) × Prof.	0.05	0.13	0.34	0.18	0.11	1.66
	Mood × Semantic (Emotive-Factive vs. Directive) × Prof.	-0.04	0.13	-0.33	-0.11	0.11	-1.01
	Mood × Task × Prof.	0.01	0.09	0.10	0.10	0.08	1.31
	Semantic (Desire vs. Directive) × Task × Prof.	0.17	0.13	1.24	-0.15	0.11	-1.36
	Semantic (Emotive-Factive vs. Directive) × Task × Prof.	-0.14	0.13	-1.10	0.10	0.11	0.89
	Mood × Semantic (Desire vs. Directive) × Task × Prof.	0.25	0.13	1.84	0.08	0.11	0.70
	Mood × Semantic (Emotive-Factive vs. Directive) × Task × Prof.	-0.19	0.13	-1.45	-0.03	0.11	-0.24
Regression-Out Probability	(Intercept)	-2.11	0.19	-11.25	-0.99	0.11	-9.15
	Mood: Subjunctive vs. Indicative	-0.15	0.22	-0.67	-0.19	0.07	-2.67
	Semantic: Desire vs. Directive	-0.16	0.20	-0.82	0.08	0.10	0.80
	Semantic: Emotive-Factive vs. Directive	0.21	0.19	1.11	0.11	0.10	1.07
	Task (Judgment vs. Comprehension)	0.22	0.14	1.59	0.42	0.07	5.82
	Proficiency (Centred)	-0.30	0.15	-2.00	-0.03	0.08	-0.35
	Sentence Length (Centred)	-0.18	0.19	-0.96	0.18	0.04	4.61
	Trial Number	0.00	0.00	-1.52	0.00	0.00	1.88
	Mood × Semantic (Desire vs. Directive)	-0.13	0.20	-0.67	-0.25	0.10	-2.43
	Mood × Semantic (Emotive-Factive vs. Directive)	-0.05	0.18	-0.28	0.10	0.10	0.96
	Mood × Task	0.14	0.14	0.99	-0.05	0.07	-0.62
	Semantic (Desire vs. Directive) × Task	-0.26	0.20	-1.32	-0.05	0.10	-0.50
	Semantic (Emotive-Factive vs. Directive) × Task	0.21	0.18	1.12	-0.01	0.10	-0.12
	Mood × Prof.	0.01	0.15	0.06	-0.07	0.08	-0.82
	Semantic (Desire vs. Directive) × Prof.	0.00	0.22	-0.01	0.09	0.12	0.80
	Semantic (Emotive-Factive vs. Directive) × Prof.	-0.13	0.20	-0.67	0.01	0.12	0.08
	Task × Prof.	-0.12	0.15	-0.79	0.06	0.08	0.69
	Mood × Semantic (Desire vs. Directive) × Task	-0.11	0.20	-0.55	0.05	0.10	0.49
	Mood × Semantic (Emotive-Factive vs. Directive) × Task	-0.03	0.18	-0.18	0.00	0.10	0.04
	Mood × Semantic (Desire vs. Directive) × Prof.	-0.32	0.22	-1.44	-0.13	0.12	-1.10
	Mood × Semantic (Emotive-Factive vs. Directive) × Prof.	0.00	0.20	0.01	-0.14	0.12	-1.25
	Mood × Task × Prof.	0.10	0.15	0.66	-0.01	0.08	-0.17
	Semantic (Desire vs. Directive) × Task × Prof.	-0.38	0.22	-1.73	0.02	0.12	0.20
	Semantic (Emotive-Factive vs. Directive) × Task × Prof.	0.45	0.20	2.23	0.02	0.12	0.20
	Mood × Semantic (Desire vs. Directive) × Task × Prof.	0.01	0.22	0.05	0.09	0.12	0.80
	Mood × Semantic (Emotive-Factive vs. Directive) × Task × Prof.	0.00	0.20	-0.02	0.03	0.12	0.28

Note: Shaded cells denote statistical significance ($p < .05$).

6.5.3 Local Measures

6.5.3.1 L1 Group

To examine the overall effect of mood and its interaction with matrix semantic property and task on reading behaviour, we calculated the following local measures: first fixation duration, gaze duration, go-past time, total reading time, regression-in probability and regression-out probability. For a more in-depth description of these measures, see Section 2.2.4.1. The fixed effect estimates for the local measures in the critical and spillover region are shown in Table 6.3.

In the critical region, a significant main effect of mood was detected for first fixation duration, gaze duration, go-past time, total reading time and regression-in probability. These reading measures were longer for indicative than subjunctive sentences, regardless of task or matrix semantic property. For total reading time, there was a significant interaction between mood, matrix semantic property (desire vs. directive) and task. Although total reading time was consistently higher for indicative than subjunctive sentences, this difference was particularly strong in the comprehension task for desire verbs, but in the judgment task for directive verbs. For regression-in probability, there was a significant interaction between mood and task. While regression-in probability was numerically higher for indicative than subjunctive sentences across tasks, this difference was more profound in the judgment task.

In the spillover region, a significant main effect of mood was observed for go-past and total reading time, such that reading times were longer for indicative than subjunctive sentences. For go-past time and total reading time, there was also a significant three-way interaction between mood, semantic (emotive-factive vs. directive) and task. Although go-past time and total reading time were consistently higher for indicative than subjunctive sentences, this pattern was modulated by task and matrix semantic property. In particular, the difference was stronger in the comprehension than the judgment task with directive verbs, but the reverse was true for emotive-factive verbs where the distinction was more pronounced in the judgment than the comprehension task.

6.5.3.2 L2 Group

To examine L2 learners' sensitivity to mood and the influence of matrix semantic property, task, and proficiency, we calculated the following local measures: first fixation duration, gaze duration, go-past time, total reading time, regression-in probability and regression-out probability. For a more in-depth description of these measures, see Section 2.2.4.1. The fixed effect estimates for the local measures in the critical and spillover region are shown in Table 6.4.

In the critical region, there was a significant main effect of mood for first fixation duration, gaze duration and go-past time, such that reading times were longer for indicative than subjunctive sentences. For first fixation duration and gaze duration, the models revealed a significant interaction between mood and task. Although these two measures were consistently longer for indicative (than subjunctive) sentences across tasks, the difference was more pronounced in the judgment task. For go-past time, there were two significant interactions: Mood × Proficiency and Mood × Semantic (Desire vs. Directive) × Task × Proficiency.[40] Figure 6.2 below shows that go-past time was typically longer for indicative than subjunctive sentences, across tasks and semantic categories, but only significantly so with emotive-factive verbs at higher proficiency levels in the comprehension task and with emotive-factive verbs at lower proficiency levels and desire verbs in the judgment task. The only exception was lower proficiency learners with directive verbs in the comprehension task.

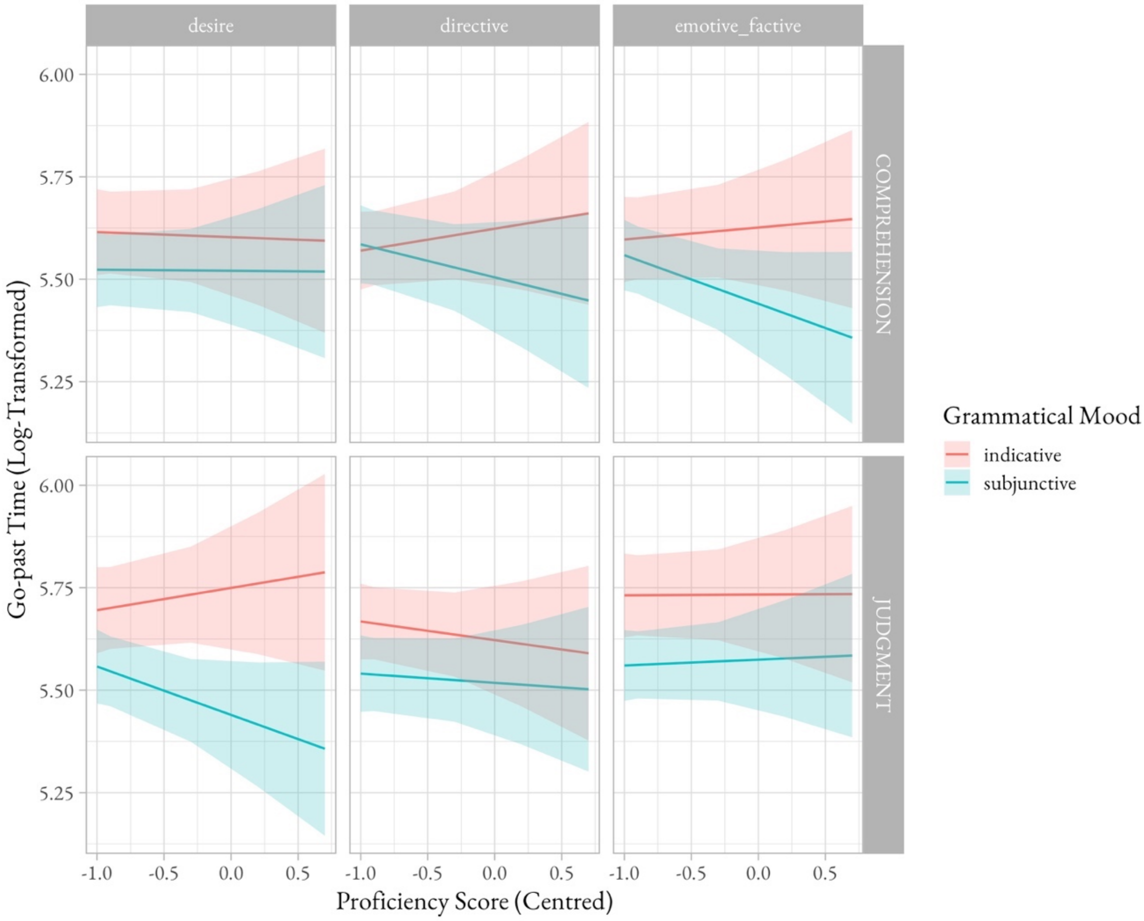


Figure 6.2 Predicted go-past time in the critical region

[40] We only explore the four-way interaction in more depth because the higher order interaction cancels out the lower order interaction.

For total reading time, there was a significant interaction between mood and proficiency. Figure 6.3 visualises this interaction. At lower proficiency levels, total reading time was identical for subjunctive and indicative sentences, but as proficiency increased, total reading time became numerically longer for indicative than subjunctive sentences.

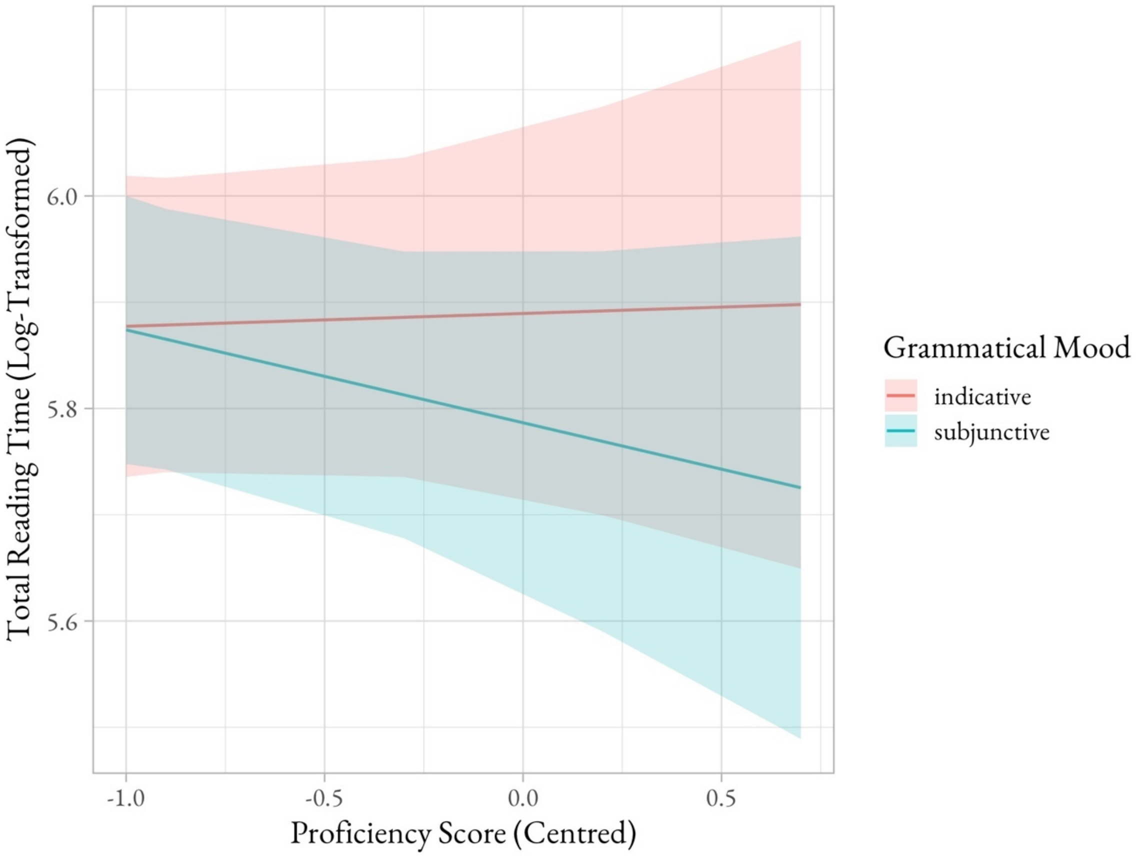


Figure 6.3 Predicted total reading time in the critical region

For regression-in probability, there was a significant interaction between mood, matrix semantic property (emotive-factive) and task. In particular, regression-in probability was, for the most part, higher for indicative than subjunctive sentences, with the exception of for emotive-factive verbs in the comprehension task and directive verbs in the judgment task.

In the spillover region, a significant main effect of mood and two-way interaction between mood and matrix semantic property (desire vs. directive) were detected for regression-out probability, with a higher regression-out probability for indicative than subjunctive sentences. In particular, regression-out probability was higher for indicative sentences with desire and directive verbs, but not with emotive-factive verbs.

There was also a significant interaction between mood and task for regression-in probability, such that regression-in probability was higher with indicative than subjunctive sentences in the comprehension task but higher with subjunctive than indicative sentences in the judgment task.

For total reading time, a significant three-way interaction was found between mood, matrix semantic property (desire vs. directive) and proficiency. Figure 6.4 shows that at lower proficiency levels, total reading time was numerically higher for indicative than subjunctive complements to desire verbs, but as proficiency increased, total reading time became numerically longer from subjunctive than indicative sentences. Conversely, at lower proficiency levels, we found a minimal difference between indicative and subjunctive complements to directive predicates, but proficiency yielded an observable, but not significant, difference in total reading time between indicative and subjunctive sentences, such that total reading time was longer for indicative sentences.

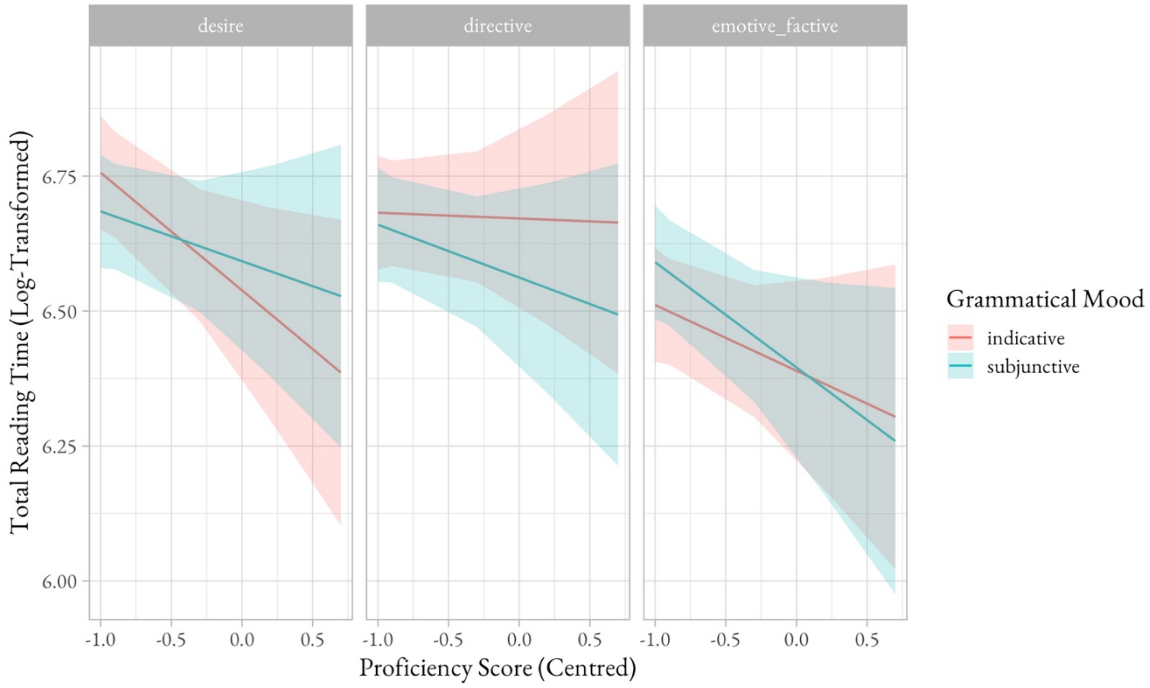


Figure 6.4 Predicted total reading time in the spillover region

6.5.4 Discussion

The data presented in Sections 6.5.2 and 6.5.3 revealed several key findings. First, native speakers of French demonstrated sensitivity to mood–modality mismatches, regardless of the lexical-semantic properties of the matrix verb or the type of secondary task used.[41] Second, L2 learners of French exhibited some sensitivity to mood–modality mismatches, but this was, in large part, influenced by a combination of factors, including the lexical-semantic properties of the matrix verb, proficiency and the type of secondary task used. The exact nature of this behaviour will be discussed in the current sub-section.

[41] Note that sometimes secondary task modulated the strength of the effect, but never the presence of an effect.

Appendix A

EXPERIMENTAL PREDICTIONS

Recall the initial hypotheses concerning the underlying causes of differences between the native speakers and the second language learners:

- Hypothesis 1: L2 learners should demonstrate increased sensitivity to mood–modality mismatches in the judgment than the comprehension task.
- Hypothesis 2: If sensitivity to mood–modality mismatches is influenced by the L1 properties, then L2 learners should demonstrate sensitivity with directive predicates first, then desire predicates and finally emotive-factive predicates at the most advanced stages of proficiency.
    - Hypothesis 2a: If we do find L1 effects, these are more likely to manifest in the comprehension task.
- Hypothesis 3: Sensitivity should improve as a function of proficiency.

Before discussing how the findings presented in the previous sections align with our initial predictions, it is important to first examine the evidence of sensitivity to mood–modality mismatches among the L2 group that was independent of task type, lexical-semantic properties and/or proficiency. In particular, we found that first fixation duration, gaze duration and go-past time in the critical region were significantly longer for indicative than subjunctive sentences. These findings indicate that L2 learners were able to detect the morphosyntactic anomaly (i.e., the ungrammaticality of the indicative) in the lower clause during real-time processing, as evidenced by first fixation duration (Liversedge, Paterson and Pickering, 1998) and gaze duration (Rayner *et al.*, 2004). It further suggests that L2 learners experienced integration difficulties having encountered the indicative form, which led to subsequent reanalysis in the form of regressions to previous parts of the sentence, as illustrated by go-past time (Rayner *et al.*, 2004). The current study clearly shows that L2 learners were able to accurately identify morphosyntactic violations during online processing, a finding consistent with previous research on L2 morphosyntactic processing (Hopp, 2006; Jackson, 2008; Sagarra and Herschensohn, 2010; Tokowicz and Warren, 2010; Foote, 2011). Although no study to date has investigated the L2 processing of the French subjunctive, our findings are in line with previous studies showing that L2 learners are able to successfully acquire the subjunctive (Howard, 2008; Bartning, Lundell and Hancock, 2012; Ayoun, 2013; McManus and Mitchell, 2015).

The findings from the current study partially confirmed Hypothesis 1 that the secondary task would influence L2 sensitivity to mood–modality mismatches. In particular, the data presented in the results section largely showed that the judgment task was the most likely of the two secondary tasks to influence sensitivity among L2 learners. For example, although L2 learners demonstrated sensitivity effects for both first fixation and gaze duration in the critical region, the effect was stronger in the judgment than the comprehension task.

However, it was rarely the secondary task alone that modulated the effect. In fact, it was often its interaction with the lexical-semantic properties of the matrix verb and/or proficiency that influenced reading behaviour. For example, the data revealed that L2 learners exhibited sensitivity to mood–modality mismatches for mean fixation duration in the judgment task, but only with directive verbs in the comprehension task. Such a finding supports not only Hypothesis 1, but also Hypothesis 2.

Further evidence in support of Hypothesis 1 was found for go-past time in the critical region. Here, most L2 learners were sensitive to mood–modality mismatches, regardless of secondary task or matrix semantic property group, with the exception of lower proficiency learners with directive verbs in the comprehension task. In other words, lower proficiency learners had difficulty engaging in reanalysis following integration difficulties with directive verbs when reading for comprehension. This finding contradicts Hypothesis 2 but supports Hypothesis 1 and Hypothesis 3. In other words, it suggests that the L1 does not always have an impact on sensitivity, but that target-deviant behaviour is more likely to surface in the comprehension than the judgment task, particularly as a function of proficiency.

Regression-in probability in the critical region was an interesting case since it both supported and contradicted Hypothesis 1 and Hypothesis 2. Here, L2 learners exhibited sensitivity to mood–modality mismatches across both tasks and matrix semantic property groups, with the exception of emotive-factive verbs in the comprehension task and directive verbs in the judgment task. The first observation (i.e., the lack of sensitivity with emotive-factive verbs in the comprehension task) was expected in light of both Hypothesis 1 and Hypothesis 2, whereas the second observation (i.e., the lack of sensitivity with directive verbs in the judgment task) was not and thus contradicts these hypotheses.

As we have shown above, the data did not always support Hypothesis 1. Further evidence against Hypothesis 1 was found for regression-in probability in the spillover region. Here, L2 learners demonstrated sensitivity in the comprehension task, but not the judgment task. It is particularly interesting to note that the only measure where sensitivity was more likely to surface in the comprehension than the judgment task was regression-in probability. Considering that several reading measures, including first fixation duration, gaze duration and go-past time, suggested target-like sensitivity to mood–modality mismatches, we must interpret this finding with caution. It is plausible that the L2 learners in question had already identified the source of the integration difficulties before leaving the interest area and did not need to perform further reanalysis.

In light of the evidence presented above and the subsequent discussion, we argue that the secondary task did indeed influence L2 sensitivity to mood–modality mismatches. More specifically, we found that L2 sensitivity was generally more likely to surface in the judgment than the comprehension task. This suggests that asking learners to make a judgment enhances their ability to detect mood–modality mismatches, which, we argue, led to (morphosyntactically) deeper processing. We must, however, recognise that it was rarely the secondary task alone, but rather its interaction with the lexical-semantic properties of the matrix verb and proficiency that modulated this effect. Nevertheless, these findings are consistent with previous studies attesting to a more target-like ability to compute morphologically marked relations in tasks that focus on form (Jackson and Bobb, 2009; Leeser, Brandl and Weissglass, 2011; Lim and Christianson, 2013b, 2015).

These findings, however, challenges the notion that the additional cognitive burden imposed by the judgment task could have hindered processing. Instead, it would appear that low-level morphological processing improved when subjected to increased cognitive burden. More generally, while our statistical models did not directly compare L1 and L2 processing patterns, these findings suggest that any L1/L2 differences in morphological processing are quantitative rather than qualitative (cf., Lim & Christianson, 2015).

An alternative interpretation for these findings could relate to the types of knowledge that L2 learners activate, namely implicit and explicit knowledge. As we discussed in the introduction to this chapter, previous research (e.g., Leeser, Brandl and Weissglass, 2011; Godfroid *et al.*, 2015) has shown that instructed L2 learners, such as those included in the current study, are more likely to rely on explicit knowledge, particularly in tasks which encourage a focus on form rather than function. The current study highlighted that while L2 learners exhibited sensitivity in both the comprehension and the judgment task, this sensitivity was stronger in the judgment task. It could therefore be argued that although L2 learners have access to both implicit and explicit knowledge of the target structure, their explicit knowledge is stronger, as evidenced by the more target-like behaviour in the judgment task, due to the high level of instructions that these learners have received in the past (cf., Leeser, Brandl and Weissglass, 2011; Godfroid *et al.*, 2015).

Although not discussed in great depth in this study, the fact that L1 sensitivity to mood-modality mismatches was occasionally stronger with certain lexical-semantic categories, most notably emotive-factive verbs, in the judgment than the comprehension task refutes, to a certain extent, the argument that stronger explicit knowledge can account for such task effects. It is widely assumed in the literature that L1 grammatical knowledge is largely implicit (VanPatten and Rothman, 2015), with the vast majority of native speakers lacking metalinguistic (explicit) awareness of specific grammatical properties. We therefore argue that deeper processing represents a more accurate explanation for the task effects observed in the current study, which would be consistent with a good enough approach to language comprehension, as espoused by Ferreira and colleagues (Christianson *et al.*, 2001; Ferreira, Bailey and Ferraro, 2002; Ferreira and Patson, 2007). According to this approach, comprehenders use both syntactic and semantic processing to parse linguistic stimuli, but sometimes not all the available morphosyntactic information is processed (e.g., when reading for comprehension) or at least not to the same degree as would be the case during tasks that encourage readers to engage in deeper (morphosyntactic) processing (e.g., when reading for judgment).

While the focus of the preceding sub-section was to investigate whether the secondary task influenced L2 sensitivity to mood–modality mismatches, it was not possible to do so without simultaneously examining the role of the lexical-semantic properties of the matrix predicate and thus L1 transfer. In particular, the discussion thus far has provided evidence both in support of and against a role for L1 transfer. Let us first briefly recap the evidence presented above.

Evidence against an L1 effect was found with go-past time in the critical region where L2 learners exhibited sensitivity to mood–modality mismatches, regardless of secondary task or matrix semantic property, with the exception of the lower proficiency learners' processing of directive verbs in the comprehension task. It must be noted, however, that this group only represented a relatively small proportion of learners in this study. It would therefore be interesting to test a wider range of lower proficiency learners to see whether this is a true effect of proficiency.

In contrast, evidence in favour of an L1 effect was observed with mean fixation duration. In particular, L2 learners exhibited sensitivity to mood–modality mismatches for mean fixation duration in the judgment task, but only with directive verbs in the comprehension task.

Some measures were inconclusive as to the role of the L1. For example, L2 learners were sensitive to mood–modality mismatches for regression-in probability in the critical region, with the exception of emotive-factive verbs in the comprehension task and directive verbs in the judgment task. As we have mentioned previously, these findings must be interpreted with caution given that several other reading measures, including first fixation duration, gaze duration and go-past time, suggested target-like sensitivity to mood-modality mismatches. It is possible that these L2 learners had already established the source of the integration difficulties, which meant further reanalysis was not necessary.

Despite the cases presented above, we did find evidence of an L1 influence and proficiency effect that were independent of the secondary task. For example, in the spillover region, there was evidence of sensitivity in both total reading time and regression-out probability, but only with directive (and desire for regression-out probability) and not emotive-factive verbs, reflecting an emerging ability to perform reanalysis following integration difficulties.

Despite a series of contradictory findings, we argue that the evidence largely suggests that the L1 did, to a certain extent, modulate the degree of sensitivity among the L2 learners tested in this study. As with task effects, L1 influence alone did not predict sensitivity, with contributing factors, such as secondary task and proficiency, also playing a role. Note that we do acknowledge the contradictory evidence but argue that it is due to noisy data (e.g., the limited number of low proficiency learners). For a more in-depth discussion of the L1 influence, see the discussion section of Study 1.

Finally, the current study largely confirms Hypothesis 3, suggesting that sensitivity improved as a function of proficiency. However, it would appear that this was often in its interaction with several contributing factors, such as secondary tasks and lexical-semantic properties. For example, as we have previously discussed, it was only the lowest proficiency learners that did not exhibit sensitivity to mood–modality mismatches for go-past time in the critical region of comprehension task. Furthermore, there was only one instance where proficiency alone predicted sensitivity to mood–modality mismatches. This was for total reading time in the critical region where only the highest proficiency learners were able to detect the morphosyntactic violation. This suggests that the ability to perform reanalysis following earlier integration difficulties develops as a function of proficiency. It is likely that at lower proficiency levels, L2 learners are able to detect the morphosyntactic violations but have trouble engaging in reanalysis following integration difficulties. In other words, increased proficiency yields a more target-like ability for reanalysis. Again, this finding is largely in line with previous research showing that knowledge of the subjunctive (Howard, 2008; Bartning, Lundell and Hancock, 2012; Ayoun, 2013; McManus and Mitchell, 2015) and, more generally, processing patterns

(McDonald, 2000, 2006; Hopp, 2006; Jackson, 2008; Foote, 2011; Lim and Christianson, 2015) improve with proficiency.

Although the analyses in Study 1 and Study 3a used the same dataset, it is important to highlight that the findings presented in both studies did not always align with each other. In particular, sensitivity effects were predominately observed in the spillover region for Study 1, but in the critical region for Study 3a. This was especially the case among the native speaker control group. The increased number of observations in Study 3a can explain this. In Study 1, the comprehension and judgment data were analysed using two separate models, whereas in Study 3a, the data from both tasks were combined. As a result, this increased the number of observations that the model could use in order to compute the estimates and ultimately the statistical power necessary to detect any significant effects and/or interactions.

In summary, the current study has revealed that L2 sensitivity to mood-modality mismatches is indeed influenced by task effects. More specifically, it would appear that learners engage in more effortful, deeper processing when required to read for judgment (as opposed to comprehension), which ultimately leads to more target-like reading patterns. It must be noted, however, that task design alone did not predict L2 sensitivity. In fact, its interaction with L1 influence and proficiency played an important role in modulating processing patterns.

Table 6.5 Mixed-effects models for L1 global (polarity subjunctive)

	Average Fixation Count			Mean Fixation Duration			Total Sentence Reading Time		
	b	*SE*	*t*	*b*	*SE*	*t*	*b*	*SE*	*t*
(Intercept)	15.68	0.85	18.39	5.41	0.02	257.47	8.21	0.13	65.32
Mood: Subjunctive vs. Indicative	-0.14	0.15	-0.94	0.00	0.00	0.75	-0.02	0.01	-1.35
Truth: True vs. False	-0.05	0.17	-0.30	-0.02	0.00	-4.26	-0.03	0.01	-2.28
Task: Judgment vs. Comprehension	2.23	0.22	10.19	0.03	0.01	5.57	0.88	0.02	49.79
Sentence Length (Centred)	0.67	0.18	3.70	-0.02	0.00	-3.95	0.04	0.01	2.76
Trial Number	-0.06	0.01	-7.29	0.00	0.00	-4.78	-0.01	0.00	-10.72
Mood × Truth	0.25	0.15	1.67	0.00	0.00	-0.22	0.01	0.01	0.89
Mood × Task	0.09	0.15	0.59	-0.01	0.00	-1.86	-0.01	0.01	-0.69
Truth × Task	0.13	0.15	0.89	0.00	0.00	-0.87	0.03	0.01	2.11
Mood × Truth × Task	0.15	0.15	0.99	0.00	0.00	0.45	0.01	0.01	0.84

Note: Shaded cells denote statistical significance ($p < 0.05$).

Table 6.6 Mixed-effects models for L2 global (polarity subjunctive)

	Mean Fixation Count			Mean Fixation Duration			Total Sentence Reading Time		
	b	*SE*	*t*	*b*	*SE*	*t*	*b*	*SE*	*t*
(Intercept)	20.76	1.57	13.22	5.45	0.03	188.60	8.56	0.17	49.73
Mood: Subjunctive vs. Indicative	-0.56	0.27	-2.03	-0.01	0.01	-1.19	-0.04	0.02	-2.63
Truth: True vs. False	-0.31	0.29	-1.08	0.00	0.01	-0.05	0.00	0.02	-0.22
Task: Judgment vs. Comprehension	5.46	0.64	8.59	0.02	0.01	1.87	0.91	0.04	24.25
Proficiency (Centred)	0.22	1.74	0.13	0.00	0.03	0.05	-0.05	0.19	-0.25
Sentence Length (Centred)	0.49	0.19	2.64	0.00	0.00	-0.75	0.03	0.01	2.50
Trial Number	-0.10	0.01	-10.96	0.00	0.00	-1.13	-0.01	0.00	-13.02
Mood × Truth	0.20	0.27	0.72	-0.01	0.01	-2.82	0.00	0.01	-0.26
Mood × Task	0.16	0.27	0.58	-0.01	0.01	-1.61	0.01	0.02	0.47
Truth × Task	0.23	0.27	0.85	0.00	0.01	0.29	0.02	0.02	1.52
Mood × Prof.	-0.48	0.31	-1.54	0.00	0.01	-0.34	-0.03	0.02	-1.96
Truth × Prof.	0.21	0.31	0.66	0.01	0.01	2.15	0.03	0.02	1.82
Task × Prof.	3.54	0.91	3.87	0.02	0.02	0.92	0.35	0.06	6.17
Mood × Truth × Task	0.06	0.27	0.21	0.00	0.01	0.30	0.01	0.02	0.47
Mood × Truth × Prof.	0.06	0.31	0.19	-0.01	0.01	-2.39	-0.01	0.02	-0.44
Mood × Task × Prof.	0.18	0.31	0.58	0.00	0.01	-0.70	0.01	0.02	0.70
Truth × Task × Prof.	0.27	0.31	0.86	0.00	0.01	0.36	0.03	0.02	1.74
Mood × Truth × Task × Prof.	0.08	0.31	0.24	0.00	0.01	0.02	-0.01	0.02	-0.62

Note: Shaded cells denote statistical significance ($p < 0.05$).

6.6 Results: Study 3b (Polarity)

6.6.1 Data Preparation and Analyses

Data preparation and analyses for Study 3b were identical to those of Study 2a and 2b. The only difference was the inclusion of task as a categorical fixed effect and the exclusion of residence abroad. Sum contrasts were coded for task to explore the difference between judgment comprehension (judgment -1, comprehension 1).

6.6.2 Global Measures

6.6.2.1 L1 Group

To examine the overall effect of mood and its interaction with truth-value and task on reading behaviour, we calculated the following global measures: mean fixation duration, mean fixation count and total sentence reading time. The fixed effect estimates for global measures are shown in Table 6.5.

Although the model did not reveal a significant main effect of mood for any of the three global measures, there was a significant main effect of truth-value for mean fixation duration and total sentence reading time, such that false sentences received longer reading times than true sentences. For all three global measures, there was a significant main effect of task. Here, reading times were consistently longer in the judgment than comprehension task. No significant interactions were observed in the L1 data.

6.6.2.2 L2 Group

To examine L2 learners' sensitivity to mood and its interaction with truth-value, task and proficiency on reading behaviour, we calculated the following global measures: mean fixation duration, mean fixation count and total sentence reading time. The fixed effect estimates for global measures are shown in Table 6.6.

The model revealed a significant main effect of mood for mean fixation count and total sentence reading time. Here, total sentence reading time and mean fixation count were higher for indicative than subjunctive sentences. For both mean fixation count and total sentence reading time, a significant main effect of task was observed, such that reading times were longer in the judgment than the comprehension task. For total sentence reading time, there was also a significant two-way interaction between mood and proficiency. Figure 6.5 visualises this interaction. Although reading times were consistently longer for indicative than subjunctive sentences, this difference increased as a function of proficiency.

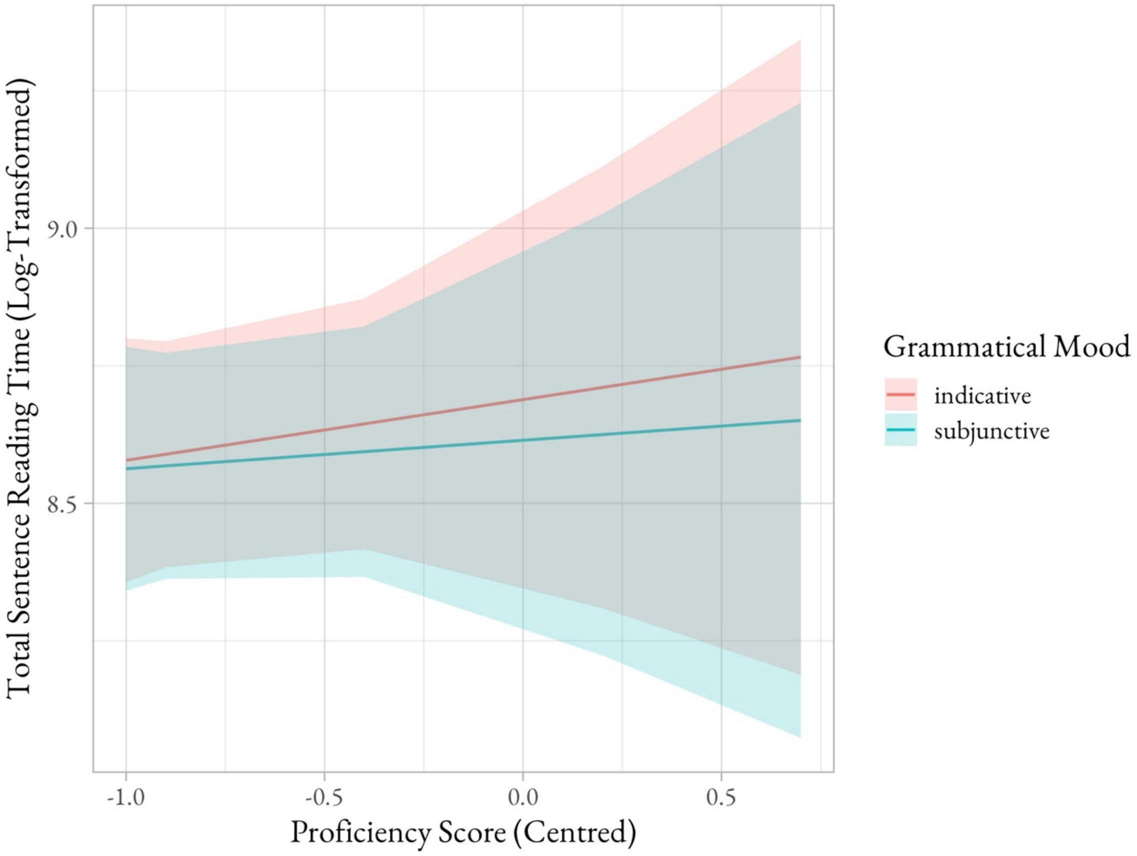


Figure 6.5 Predicted total sentence reading time

Finally, for mean fixation duration, there were two significant two-way interactions between mood and truth-value, and between truth-value and proficiency, as well as a significant three-way interaction between mood, truth-value and proficiency. Figure 6.6 explores the three-way interaction further. It shows that Indicative-False sentences received numerically longer mean fixation durations than Subjunctive-False sentences at the lower end of the proficiency spectrum. However, this trend reversed as proficiency increased to the extent that at the higher end of the proficiency spectrum, Subjunctive-False sentences received numerically longer mean fixation durations than Indicative-False sentences. Conversely, Indicative-True sentences attracted numerically longer mean fixation durations than subjunctive sentences particularly as a function of proficiency.

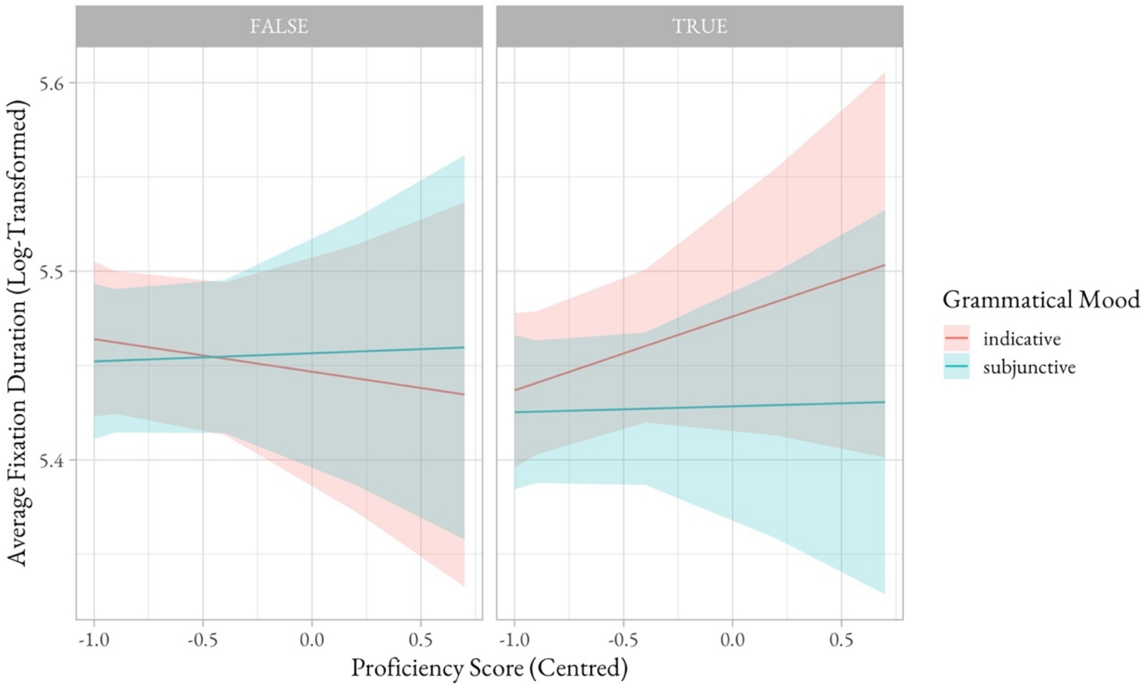


Figure 6.6 Predicted mean fixation duration

Table 6.7 Mixed-effects models for L1 local (polarity subjunctive)

		Critical Region			Spillover Region		
		b	*SE*	*t*/*z*	*b*	*SE*	*t*/*z*
First Fixation Duration	(Intercept)	5.27	0.04	139.62	5.36	0.02	220.61
	Mood: Subjunctive vs. Indicative	0.00	0.02	-0.24	-0.02	0.01	-2.39
	Truth: True vs. False	0.00	0.01	0.22	-0.01	0.01	-1.86
	Task: Judgment vs. Comprehension	0.04	0.01	3.19	0.00	0.01	0.25
	Sentence Length (Centred)	-0.01	0.02	-0.30	-0.01	0.01	-0.70
	Trial Number	0.00	0.00	1.00	0.00	0.00	-1.66
	Mood × Truth	-0.01	0.01	-0.73	0.01	0.01	1.25
	Mood × Task	-0.01	0.01	-0.68	-0.01	0.01	-1.17
	Truth × Task	0.02	0.01	1.35	0.00	0.01	0.69
	Mood × Truth × Task	0.00	0.01	0.36	-0.01	0.01	-0.98
Gaze Duration	(Intercept)	5.29	0.04	131.52	5.59	0.03	163.38
	Mood: Subjunctive vs. Indicative	-0.01	0.02	-0.30	-0.04	0.01	-4.22
	Truth: True vs. False	0.00	0.01	-0.29	-0.03	0.01	-3.54
	Task: Judgment vs. Comprehension	0.04	0.01	3.41	-0.02	0.01	-1.76
	Sentence Length (Centred)	-0.01	0.02	-0.43	0.15	0.02	9.48
	Trial Number	0.00	0.00	1.07	0.00	0.00	-1.87
	Mood × Truth	-0.01	0.01	-0.85	0.02	0.01	1.56
	Mood × Task	-0.01	0.01	-1.03	-0.02	0.01	-1.66
	Truth × Task	0.02	0.01	1.80	0.00	0.01	0.51
	Mood × Truth × Task	0.00	0.01	0.38	0.00	0.01	0.15
Go-Past Time	(Intercept)	5.32	0.04	125.43	5.71	0.04	154.76
	Mood: Subjunctive vs. Indicative	0.01	0.02	0.28	-0.05	0.01	-5.42
	Truth: True vs. False	0.01	0.01	0.48	-0.04	0.01	-4.29
	Task: Judgment vs. Comprehension	0.04	0.01	2.80	0.00	0.01	0.33
	Sentence Length (Centred)	0.00	0.02	0.01	0.21	0.02	11.23
	Trial Number	0.00	0.00	0.92	0.00	0.00	-2.04
	Mood × Truth	-0.01	0.01	-0.47	0.01	0.01	0.68
	Mood × Task	-0.01	0.01	-1.13	-0.03	0.01	-3.38
	Truth × Task	0.02	0.01	1.21	0.01	0.01	0.96
	Mood × Truth × Task	0.01	0.01	1.09	0.01	0.01	1.15
Total Reading Time	(Intercept)	5.54	0.06	86.85	6.04	0.06	107.59
	Mood: Subjunctive vs. Indicative	0.03	0.03	1.11	-0.06	0.01	-5.25
	Truth: True vs. False	-0.02	0.02	-1.06	-0.04	0.01	-3.67
	Task: Judgment vs. Comprehension	0.10	0.02	5.89	0.07	0.01	5.39
	Sentence Length (Centred)	0.00	0.03	-0.14	0.24	0.02	12.33
	Trial Number	0.00	0.00	-0.98	0.00	0.00	-4.72
	Mood × Truth	-0.01	0.02	-0.45	-0.01	0.01	-1.03
	Mood × Task	-0.01	0.02	-0.82	-0.04	0.01	-2.93
	Truth × Task	0.01	0.02	0.46	0.01	0.01	1.08
	Mood × Truth × Task	0.01	0.02	0.69	0.00	0.01	0.28
Regression-In Probability	(Intercept)	-0.51	0.25	-2.04	-1.48	0.15	-10.03
	Mood: Subjunctive vs. Indicative	-0.52	0.11	-4.57	-0.17	0.07	-2.47
	Truth: True vs. False	-0.07	0.08	-0.88	-0.01	0.07	-0.16
	Task: Judgment vs. Comprehension	0.10	0.08	1.29	0.15	0.07	2.25
	Sentence Length (Centred)	-0.15	0.12	-1.29	0.22	0.06	3.35
	Trial Number	-0.01	0.00	-2.30	-0.01	0.00	-2.07
	Mood × Truth	-0.05	0.08	-0.70	-0.12	0.07	-1.74
	Mood × Task	-0.07	0.08	-0.97	-0.08	0.07	-1.19
	Truth × Task	0.19	0.08	2.45	0.04	0.07	0.65
	Mood × Truth × Task	-0.01	0.08	-0.11	-0.02	0.07	-0.30
Regression-Out Probability	(Intercept)	-3.13	0.36	-8.80	-1.93	0.15	-12.48
	Mood: Subjunctive vs. Indicative	0.48	0.27	1.78	-0.01	0.07	-0.18
	Truth: True vs. False	0.18	0.17	1.04	-0.01	0.07	-0.14
	Task: Judgment vs. Comprehension	0.23	0.17	1.35	0.00	0.07	-0.06
	Sentence Length (Centred)	0.29	0.23	1.26	0.19	0.06	3.00
	Trial Number	0.00	0.01	0.20	0.01	0.00	1.70
	Mood × Truth	0.18	0.17	1.03	-0.03	0.07	-0.51

	Critical Region			Spillover Region		
	b	*SE*	*t*/*z*	*b*	*SE*	*t*/*z*
Mood × Task	-0.27	0.17	-1.58	-0.04	0.07	-0.57
Truth × Task	0.01	0.17	0.06	0.08	0.07	1.25
Mood × Truth × Task	0.09	0.17	0.52	0.08	0.07	1.22

Note: Shaded cells denote statistical significance ($p < .05$).

Table 6.8 Mixed-effects models for L2 local (polarity subjunctive)

		Critical Region			Spillover Region		
		b	*SE*	*t/z*	*b*	*SE*	*t/z*
First Fixation Duration	(Intercept)	5.47	0.04	140.89	5.39	0.03	174.08
	Mood: Subjunctive vs. Indicative	-0.03	0.02	-1.83	0.00	0.01	0.43
	Truth: True vs. False	0.03	0.01	1.81	-0.01	0.01	-0.68
	Task: Judgment vs. Comprehension	0.03	0.01	1.77	0.02	0.01	1.66
	Proficiency (Centred)	0.00	0.04	0.06	-0.01	0.03	-0.43
	Sentence Length (Centred)	-0.01	0.01	-0.85	0.00	0.01	0.33
	Trial Number	0.00	0.00	-2.63	0.00	0.00	1.66
	Mood × Truth	-0.02	0.01	-1.34	-0.01	0.01	-0.77
	Mood × Task	-0.03	0.01	-2.16	0.00	0.01	-0.32
	Truth × Task	0.01	0.01	0.60	-0.01	0.01	-0.83
	Mood × Prof.	-0.03	0.02	-1.49	0.00	0.01	-0.18
	Truth × Prof.	0.02	0.02	1.27	0.01	0.01	1.05
	Task × Prof.	-0.01	0.02	-0.76	0.01	0.01	0.65
	Mood × Truth × Task	-0.02	0.01	-1.12	0.02	0.01	1.36
	Mood × Truth × Prof.	-0.01	0.02	-0.51	-0.01	0.01	-0.81
	Mood × Task × Prof.	-0.01	0.02	-0.73	0.01	0.01	0.51
	Truth × Task × Prof.	0.01	0.02	0.61	-0.01	0.01	-0.80
	Mood × Truth × Task × Prof.	-0.01	0.02	-0.58	0.02	0.01	1.42
Gaze Duration	(Intercept)	5.50	0.04	125.80	5.67	0.05	110.45
	Mood: Subjunctive vs. Indicative	-0.02	0.02	-1.27	0.00	0.02	-0.12
	Truth: True vs. False	0.02	0.02	1.55	-0.04	0.02	-2.31
	Task: Judgment vs. Comprehension	0.03	0.02	2.23	-0.01	0.02	-0.33
	Proficiency (Centred)	0.00	0.05	-0.06	-0.06	0.05	-1.21
	Sentence Length (Centred)	0.00	0.01	-0.15	0.19	0.02	8.11
	Trial Number	0.00	0.00	-2.80	0.00	0.00	1.26
	Mood × Truth	-0.02	0.02	-1.05	-0.01	0.02	-0.51
	Mood × Task	-0.04	0.02	-2.72	0.01	0.02	0.63
	Truth × Task	0.00	0.02	0.26	-0.01	0.02	-0.38
	Mood × Prof.	-0.03	0.02	-1.69	0.00	0.02	-0.10
	Truth × Prof.	0.02	0.02	0.96	0.00	0.02	0.08
	Task × Prof.	-0.01	0.02	-0.60	0.02	0.02	1.05
	Mood × Truth × Task	-0.02	0.02	-1.11	0.00	0.02	-0.02
	Mood × Truth × Prof.	-0.01	0.02	-0.45	-0.01	0.02	-0.38
	Mood × Task × Prof.	-0.01	0.02	-0.70	0.02	0.02	1.22
	Truth × Task × Prof.	0.01	0.02	0.34	-0.03	0.02	-1.35
	Mood × Truth × Task × Prof.	-0.01	0.02	-0.35	0.01	0.02	0.62
Go-Past Time	(Intercept)	5.53	0.05	120.90	5.88	0.06	99.01
	Mood: Subjunctive vs. Indicative	-0.01	0.02	-0.37	0.00	0.02	0.20
	Truth: True vs. False	0.02	0.02	0.98	-0.03	0.02	-2.15
	Task: Judgment vs. Comprehension	0.05	0.02	2.92	0.01	0.02	0.50
	Proficiency (Centred)	-0.01	0.05	-0.22	-0.08	0.06	-1.37
	Sentence Length (Centred)	0.00	0.01	0.12	0.26	0.03	9.02
	Trial Number	0.00	0.00	-2.30	0.00	0.00	-0.40
	Mood × Truth	-0.02	0.02	-1.50	0.00	0.02	0.24
	Mood × Task	-0.04	0.02	-2.44	-0.02	0.02	-1.44
	Truth × Task	0.00	0.02	-0.07	0.01	0.02	0.52
	Mood × Prof.	-0.01	0.02	-0.70	-0.01	0.02	-0.72
	Truth × Prof.	0.01	0.02	0.73	0.03	0.02	1.52
	Task × Prof.	0.00	0.02	-0.09	0.02	0.02	0.89
	Mood × Truth × Task	-0.02	0.02	-1.42	-0.01	0.02	-0.93
	Mood × Truth × Prof.	-0.02	0.02	-0.96	0.02	0.02	1.31
	Mood × Task × Prof.	-0.02	0.02	-0.86	-0.02	0.02	-0.84
	Truth × Task × Prof.	0.00	0.02	-0.20	-0.02	0.02	-1.18
	Mood × Truth × Task × Prof.	-0.01	0.02	-0.42	-0.02	0.02	-1.35
Total Reading Time	(Intercept)	5.81	0.07	85.73	6.34	0.08	79.75
	Mood: Subjunctive vs. Indicative	0.04	0.03	1.65	-0.03	0.02	-1.44
	Truth: True vs. False	-0.01	0.02	-0.51	-0.03	0.02	-1.71

		Critical Region			Spillover Region		
		b	*SE*	*t/z*	*b*	*SE*	*t/z*
	Task: Judgment vs. Comprehension	0.10	0.02	4.59	0.07	0.02	3.87
	Proficiency (Centred)	-0.05	0.07	-0.76	-0.11	0.08	-1.35
	Sentence Length (Centred)	0.04	0.02	2.02	0.27	0.03	9.11
	Trial Number	0.00	0.00	-3.76	0.00	0.00	-5.97
	Mood × Truth	-0.02	0.02	-1.05	0.01	0.02	0.67
	Mood × Task	-0.05	0.02	-2.31	0.00	0.02	0.00
	Truth × Task	-0.01	0.02	-0.66	0.04	0.02	1.99
	Mood × Prof.	-0.04	0.03	-1.59	-0.05	0.02	-2.33
	Truth × Prof.	0.01	0.03	0.37	0.04	0.02	1.97
	Task × Prof.	-0.01	0.03	-0.38	0.01	0.02	0.36
	Mood × Truth × Task	-0.02	0.02	-0.98	0.00	0.02	-0.02
	Mood × Truth × Prof.	0.00	0.03	-0.12	0.03	0.02	1.55
	Mood × Task × Prof.	-0.01	0.03	-0.46	0.02	0.02	0.80
	Truth × Task × Prof.	-0.01	0.03	-0.54	0.00	0.02	0.19
	Mood × Truth × Task × Prof.	0.01	0.03	0.45	0.00	0.02	0.15
Regression-In Probability	(Intercept)	-0.65	0.13	-4.88	-1.07	0.12	-8.64
	Mood: Subjunctive vs. Indicative	-0.12	0.10	-1.18	-0.06	0.08	-0.71
	Truth: True vs. False	0.00	0.09	-0.05	0.08	0.08	0.95
	Task: Judgment vs. Comprehension	-0.14	0.09	-1.52	-0.03	0.08	-0.33
	Proficiency (Centred)	-0.24	0.10	-2.38	-0.20	0.10	-2.11
	Sentence Length (Centred)	0.07	0.07	1.03	0.21	0.04	4.65
	Trial Number	0.00	0.00	-1.09	-0.01	0.00	-3.03
	Mood × Truth	0.04	0.09	0.47	-0.03	0.08	-0.34
	Mood × Task	-0.10	0.09	-1.12	0.17	0.08	1.99
	Truth × Task	-0.04	0.09	-0.45	0.15	0.08	1.74
	Mood × Prof.	-0.24	0.10	-2.35	-0.07	0.10	-0.69
	Truth × Prof.	0.07	0.10	0.71	0.09	0.10	0.93
	Task × Prof.	-0.15	0.10	-1.48	-0.13	0.10	-1.41
	Mood × Truth × Task	-0.05	0.09	-0.55	-0.16	0.08	-1.87
	Mood × Truth × Prof.	0.15	0.10	1.46	-0.06	0.10	-0.60
	Mood × Task × Prof.	0.03	0.10	0.25	0.23	0.10	2.43
	Truth × Task × Prof.	-0.08	0.10	-0.75	0.05	0.10	0.50
	Mood × Truth × Task × Prof.	-0.07	0.10	-0.66	-0.16	0.10	-1.67
Regression-Out Probability	(Intercept)	-2.62	0.24	-10.85	-1.08	0.13	-8.49
	Mood: Subjunctive vs. Indicative	0.06	0.19	0.33	-0.02	0.08	-0.20
	Truth: True vs. False	-0.16	0.17	-0.96	0.02	0.08	0.24
	Task: Judgment vs. Comprehension	0.29	0.17	1.72	-0.09	0.08	-1.03
	Proficiency (Centred)	-0.11	0.19	-0.61	0.06	0.10	0.66
	Sentence Length (Centred)	0.01	0.12	0.09	0.13	0.05	2.78
	Trial Number	0.00	0.00	0.38	-0.01	0.00	-2.48
	Mood × Truth	-0.04	0.17	-0.24	0.17	0.08	2.00
	Mood × Task	0.10	0.17	0.61	-0.19	0.08	-2.21
	Truth × Task	0.08	0.17	0.46	0.09	0.08	1.04
	Mood × Prof.	0.12	0.19	0.64	-0.06	0.10	-0.58
	Truth × Prof.	-0.21	0.19	-1.15	0.03	0.10	0.29
	Task × Prof.	0.11	0.19	0.57	0.00	0.10	0.00
	Mood × Truth × Task	0.05	0.17	0.32	0.05	0.08	0.63
	Mood × Truth × Prof.	-0.11	0.19	-0.59	0.21	0.10	2.19
	Mood × Task × Prof.	0.12	0.19	0.64	-0.11	0.10	-1.12
	Truth × Task × Prof.	0.04	0.19	0.23	0.07	0.10	0.72
	Mood × Truth × Task × Prof.	0.03	0.19	0.15	-0.04	0.10	-0.41

Note: Shaded cells denote statistical significance ($p < 0.05$).

6.6.3 Local Measures

6.6.3.1 L1 Group

To examine the overall effect of mood and its interaction with truth-value and task on reading behaviour, we calculated the following local measures: first fixation duration, gaze duration, go-past time, total reading time, regression-in probability and regression-out probability. For a more in-depth description of these measures, see Section 2.2.4.1. The fixed effect estimates for the local measures in the critical and spillover regions are shown in Table 6.7.

In the critical region, the model revealed a significant main effect of mood as well as a significant two-way interaction between truth-value and task for regression-in probability. These probabilities were higher for false than true sentences in the comprehension task, but higher for true than false sentences in the judgment task. Furthermore, a significant main effect of task was detected for first fixation duration, gaze duration, go-past time and total reading time. In particular, these measures were higher in the judgment than the comprehension task.

In the spillover region, a significant main effect of mood was found for first fixation duration, gaze duration, go-past time, total reading time, regression-in probability. These measures were higher with indicative than subjunctive sentences. There was a significant main effect of truth-value for gaze duration, go-past time and total reading time, such that these measures were higher for false than true sentences. The model further revealed a significant main effect of task for total reading time and regression-in probability. These measures were higher in the judgment than the comprehension task.

A significant two-way interaction was detected between mood and task for go-past time and total reading time. In particular, we found that go-past time and total reading time were longer for subjunctive than indicative sentences in the comprehension task, but longer for indicative than subjunctive sentences in the judgment task.

6.6.3.2 L2 Group

To examine L2 learners' sensitivity to mood and the influence of truth-value, task, and proficiency on reading behaviour, we calculated the following local measures: first fixation duration, gaze duration, go-past time, total reading time, regression-in probability and regression-out probability. For a more in-depth description of these measures, see Section 2.2.4.1. The fixed effect estimates for the local measures in the critical and spillover regions are shown in Table 6.8.

In the critical region, none of the local measures exhibited a significant main effect of mood or truth-value. For gaze duration, go-past time and total reading time, there was a significant main effect of task as well as a significant interaction between mood and task. In the comprehension task, all three measures were longer for subjunctive than indicative sentences. However, in the judgment task, although total reading time was higher for subjunctive than indicative sentences, gaze duration and go-past time were higher for indicative than subjunctive sentences.

In the spillover region, we did not find a significant main effect of mood for any of the local measures. However, there was a significant main effect of truth-value for gaze duration and go-past time, such that these measures were higher for false than true sentences. The model further revealed a significant main effect of task as well as a significant two-way interaction between mood and proficiency for total reading time. Figure 6.7 visualises this interaction. At the lowest proficiency levels tested, total reading time was numerically longer for subjunctive than indicative sentences, but this trend reversed as a function of proficiency, such that total reading time was numerically longer for indicative than subjunctive sentences at higher proficiency levels.

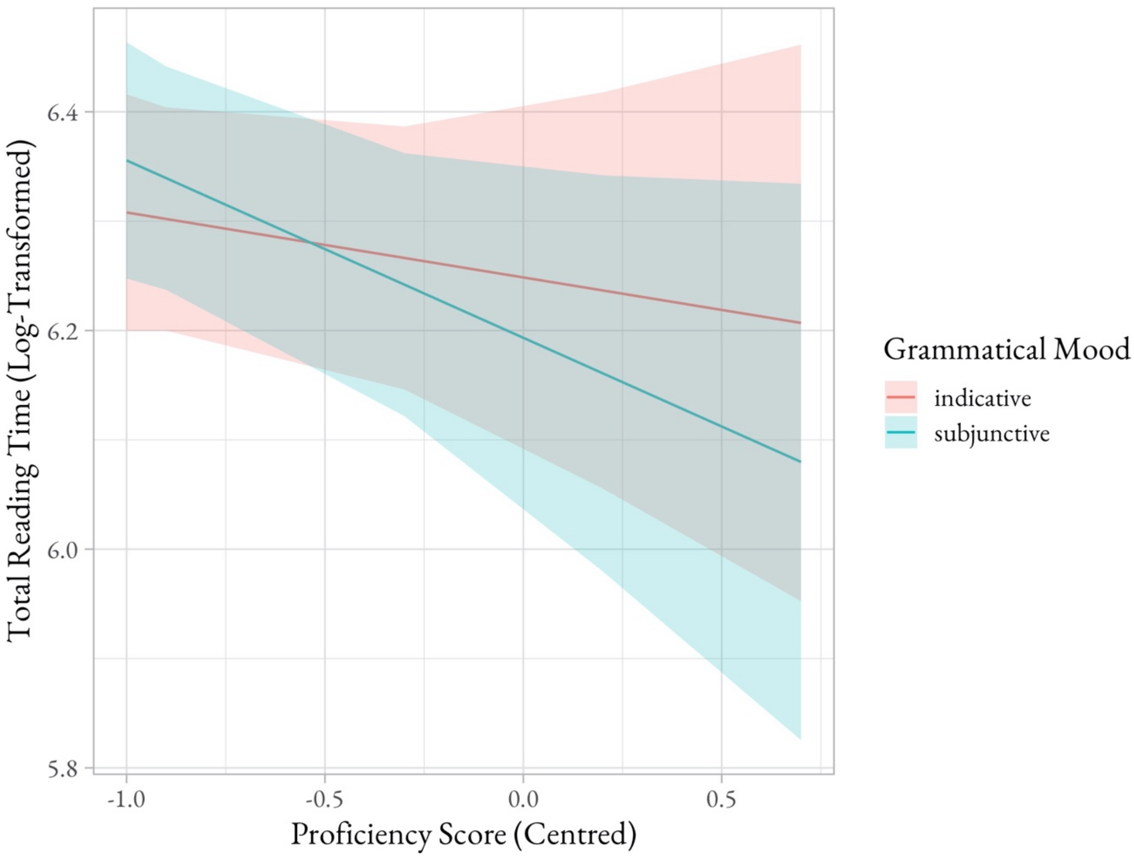


Figure 6.7 Predicted total reading time in the spillover region

For regression-in probability, there was a significant three-way interaction between mood, task and proficiency for regression-in probability. Figure 6.8 visualises this interaction. In the comprehension task, regression-in probability was numereically higher for subjunctive than indicative sentences at lower proficiency levels, but as proficiency increased, this trend reversed to the extent that regression-in probability was significantly higher for indicative than subjunctive sentences at higher proficiency levels. The opposite pattern was found in the judgment data. At lower proficiency levels, regression-in probability was numerically higher for indicative than subjunctive sentences, but this trend reversed as proficiency increased, to the extent that regression-in probability was numerically higher for subjunctive than indicative sentences.

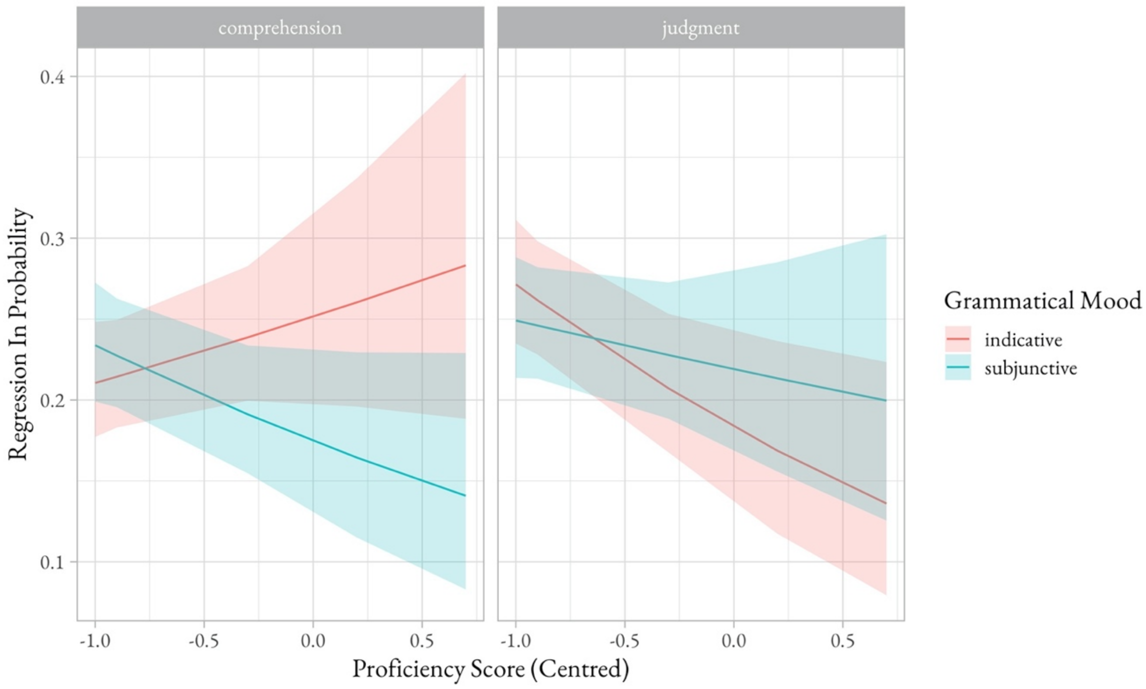


Figure 6.8 Predicted regression-in probability in the spillover region

For regression-out probability, there was a significant two-way interaction between mood and task. In the comprehension task, regression-out probability was higher for subjunctive than indicative sentences, whereas in the judgment task, regression-out probability was higher for indicative than subjunctive sentences. The model further revealed a significant two-way interaction between mood and truth-value and a significant three-way interaction between mood, truth-value and proficiency. Figure 6.9 visualises this interaction. It shows that regression-out probability was numerically higher for Subjunctive-False than Indicative-False sentences at the lowest proficiency levels tested, but this trend reversed a function of proficiency to the extent that regression-out probability was significantly higher for indicative than subjunctive sentences at the highest proficiency levels. In contrast, at the lowest proficiency levels tested, regression-out probability was almost identical for Indicative-True and Subjunctive-True sentences. However, increased proficiency yielded an observable, but not significant, difference between indicative and subjunctive sentences, such that regression-out probability was higher for subjunctive than indicative sentences at the highest proficiency levels tested.

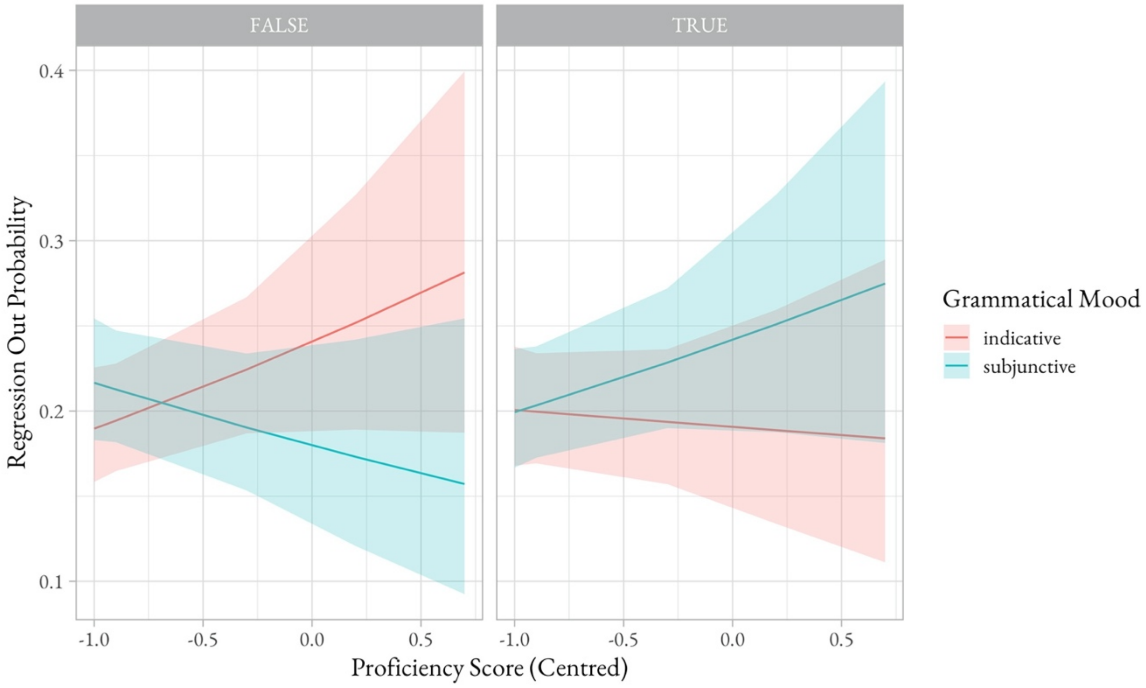


Figure 6.9 Predicted regression-out probability in the spillover region

6.6.4 Discussion

The overarching aim of the current study was to investigate whether the type of secondary task influenced sensitivity to discourse-pragmatic constraints in polarity contexts among both L1 and L2 speakers. The data presented in Sections 6.6.2 and 6.6.3 revealed several key findings both in terms of the L1 and L2 behaviour. Unlike Study 3a (obligatory subjunctives) where native speakers exhibited sensitivity to mood–modality mismatches during real-time processing, regardless of task type or the lexical-semantic properties of the matrix verb, this was not always the case in the current study. In fact, we found no evidence of sensitivity to the discourse-pragmatic constraints among the L1 or the L2 group, only a distinct preference for the subjunctive among L1 speakers and a certain level of indeterminacy between the indicative and the subjunctive among the L2 speakers. The exact nature of this behaviour will be discussed in further detail below.

Appendix A

EXPERIMENTAL PREDICTIONS

Recall the initial hypotheses concerning the underlying causes of differences between the native speakers and the second language learners:

- Hypothesis 1: L2 learners should demonstrate increased sensitivity to the discourse-pragmatic constraints in the judgment than the comprehension task.
- Hypothesis 2: Sensitivity should improve as a function of proficiency.

When formulating these hypotheses, we did not anticipate that the native speakers would not exhibit any sensitivity to the discourse-pragmatic constraints in polarity contexts and as such did not speculate on the possible causes for such behaviour. However, it is undeniable that the native speaker behaviour has important implications for L2 acquisition. Put simply, you cannot expect the L2 group to behave in a target-like, consistent manner when the control group does not. For this reason, the current discussion will be divided into two parts. The first will examine the possible causes for this absence of sensitivity. The second will explore the L2 behaviour. The discussion here will be guided, in large part, by our initial hypotheses concerning the underlying causes of differences between the L1 and L2 speakers, as presented above. Given the overarching aim of the current study, both sections will focus in particular on the influence of the secondary task on sensitivity.

L1 DATA

As discussed in the introduction, the native speaker behaviour was not fully in line with our expectations. In particular, we found that the L1 speakers were not sensitive to the discourse-pragmatic constraints in polarity contexts. Recall that although the subjunctive is typically the expected form in the embedded complement clauses of negated epistemics, the indicative can also be used depending on the speaker's (or the subject of the matrix clause's) commitment to the truth-value of the embedded clause. If the speaker (or the subject of the matrix clause) is committed to the truth-value of the embedded clause, then the indicative is licensed.

The current study indicates that native speakers largely considered the subjunctive to be the form *par excellence* under negated epistemics in French, regardless of their commitment to the truth-value of the embedded clause. This preference was observed in several measures, including first fixation duration, gaze duration, go-past time, total reading time and regression-in probability in the spillover region, where reading times (or probabilities) were higher for indicative than subjunctive sentences. Note, however, that go-past time and total reading time were longer for subjunctive than indicative sentences in the comprehension task, but not the judgment task. The influence of the secondary task will be discussed in more depth in the latter part of this section. Nevertheless, the presence of sensitivity effects in both the early and late measures suggests that native speakers were able to both detect the morphosyntactic violations and engage in reanalysis following earlier integration difficulties (Rayner *et al.*, 2004).

Despite a lack of sensitivity to the discourse-pragmatic constraints, the current study did reveal that sensitivity, at least to the ungrammaticality of the indicative, was more likely to be detected in the judgment task. For example, in the spillover region of the comprehension task, go-past time and total reading time were longer for subjunctive than indicative sentences, whereas the reverse was true for the judgment task. This is a prime example of where sensitivity effects were absent in the comprehension task. The presence of sensitivity effects for go-past time and total reading time in the judgment task, but their absence in the comprehension task, is particularly revealing given that both go-past time and total reading time have been shown to index higher-order processing, such as reanalysis (Liversedge, Paterson and Pickering, 1998; Rayner *et al.*, 2004). In particular, it suggests that the judgment task led to more effortful, deeper processing even among the native speakers.

Based on the findings reported above, our dataset largely indicates that native speakers of French prefer the subjunctive over the indicative, regardless of the discourse-pragmatic conditions. For a more in-depth discussion of possible explanations, see Section 5.10 (the discussion section of Study 2).

However, it is important to recall that the overarching aim of the current study was to investigate whether the type of secondary task influenced sensitivity to discourse-pragmatic constraints. We argue that the current study largely supports the view that sensitivity is more likely to be detected in the judgment task than the comprehension task. Initially, we did not anticipate that native speakers would demonstrate different degrees of sensitivity depending on the secondary task, especially as this was not the case with mood use in obligatory contexts. There are several reasons why this may be the case.

It is widely acknowledged that the polarity subjunctive is a highly infrequent structure in the input (Poplack, 1989, 1990; O'Connor DiVito, 1997; Soutet, 2000; Poplack, Lealess and Dion, 2013), to the extent that some scholars have even argued that the subjunctive in such contexts is purely a marker of formality (Müller and Elsass, 1985; O'Connor DiVito, 1997; Lenoble-Pinson and Grevisse, 2009; Jeppesen Kragh, 2010). In psycholinguistic terms, this means that the grammatical features associated with the polarity subjunctive are activated relatively infrequently and as such, the activation threshold levels for this structure remain low

(Yang, 2018). As we have discussed throughout this chapter, tasks that encourage readers to focus on morphosyntactic form are more likely to result in more effortful, deeper processing, a type of processing that is required in order to retrieve and activate features with low activation thresholder levels and thus infrequent structures (Christianson *et al.*, 2001; Ferreira, Bailey and Ferraro, 2002; Ferreira and Patson, 2007).

The more target-like behaviour in the judgment task could also be explained in sociolinguistic terms. In the previous paragraph, we discussed how the focus on form in the judgment task led to more effortful, deeper processing and that this enabled L2 learners to more easily retrieve features with low activation threshold levels from the mental lexicon. However, as we have already highlighted, the subjunctive, particularly in French, is often subject to socio-stylistic variation, to the extent that it is typically considered a marker of formality, predominately found in written language (Müller and Elsass, 1985; O'Connor DiVito, 1997; Lenoble-Pinson and Grevisse, 2009; Jeppesen Kragh, 2010). Given that the ideology of the standard language exerts strong normative pressure on (French) language (Lodge, 1993), it is possible that tasks requiring a focus on form inevitably lead native speakers to monitor their language use more closely in order to align it with the written form.

Since very little research has empirically tested the processing and knowledge of the polarity subjunctive in French, it is difficult to determine with any certainty whether the variability in the L1 data is the result of (extra-)linguistic factors, or whether it is simply the by-product of the experimental design. In order to examine this, it would be advisable to employ an alternative experimental design, an example of which would be the Iverson *et al.* (2008) study that provided short stories in order to situate the test sentence within the wider discourse-pragmatic context. Despite its focus on L2 Spanish, a replication of this study in French would be particularly interesting since the L1 Spanish speakers did not exhibit the same level of variability as the L1 French speakers did in our study.

L2 DATA

Turning now to the L2 data, as we mentioned in the introduction, the current study revealed several key findings. In particular, the L2 speakers did not exhibit any sensitivity to the discourse-pragmatic constraints of the polarity subjunctive but did demonstrate a certain level of indeterminacy between the indicative and the subjunctive, which was constrained, to a certain extent, by the secondary task, but not by proficiency. The current sub-section discusses these findings in more depth.

As was the case with the L1 speakers, the L2 speakers exhibited a (distinct) preference for one particular mood, a preference that was largely independent of truth-value. However, unlike the L1 speakers' preference for the subjunctive mood, the L2 speakers' preferred mood largely depended on the level and location of the reading measure, that is, whether it was a global measure or a local measure in the critical or spillover region.

In terms of the global measures and the local measures in the spillover region, L2 speakers largely mirrored the L1 speakers with respect to their preference for the subjunctive over the indicative. For example, in two of the global measures (total sentence reading time and mean fixation duration), L2 learners exhibited a distinct preference for the subjunctive over the indicative, such that both total sentence reading time and mean fixation duration were higher for indicative than subjunctive sentences. In the spillover region, we observed a similar preference for the subjunctive in total reading time at the highest proficiency levels, regression-in probability in the comprehension task and regression-out probability in the judgment task, suggesting an emerging ability for reanalysis (Rayner and Pollatsek, 1989). The two latter measures are particularly interesting since L2 speakers favoured the indicative with regression-in probability in the judgment task and regression-out probability in the comprehension task.

This preference for the indicative among the L2 group was further reflected in gaze duration, go-past time and total reading time in the critical region, with the exception of gaze duration and go-past time in the judgment task. For these measures, reading times were longer for subjunctive than indicative sentences. This indicates that the indicative was considered to be the grammatical form in the interlanguage L2 grammar. This may indeed be the by-product of L1 transfer since negated epistemics in English only allow indicative complement phrases and would largely be in line with a Feature Reassembly account of language acquisition (Lardiere, 2009), as discussed in Section 5.3.

It is particularly interesting to note that L2 speakers mostly exhibited a preference for the subjunctive in the global measures and in the spillover region, but a preference for the indicative in the critical region. As we have previously discussed, delayed sensitivity is typical among the L2 population (Felser and Cunnings, 2012; Felser *et al.*, 2012; Lim and Christianson, 2015; Boxell and Felser, 2017).

Given this optionality between the indicative and the subjunctive, it is not surprising that there was an absence of sensitivity to the discourse-pragmatic constraints. Turning now to the main focus of the current study: the influence of the secondary task. As with the L1 speakers, the dataset largely suggests that sensitivity was more likely to surface in the judgment than the comprehension task despite a handful of counterexamples.

On one hand, the data showed that sensitivity was more likely to be observed in judgment than comprehension. For example, gaze duration and go-past time were longer for indicative than subjunctive sentences in the critical region of the judgment task, but not in the comprehension task. A similar finding was observed for regression-out probability in the spillover region. Although such a finding does not necessarily reflect sensitivity to the discourse-pragmatic constraints, it does suggest that L2 learners are aware, to a certain extent, that the indicative is ungrammatical in certain polarity contexts. This awareness, however, was only observed in the judgment task. Given that these measures have been shown to index various stages of processing, including sensitivity to morphosyntactic violations and reanalysis following integration difficulties, this indicates that the judgment task led to more effortful, deeper processing among the L2 learners.

The logic of this argument, however, is challenged to a certain extent when you examine the handful of cases where target-like behaviour was more likely to surface in the comprehension than the judgment task. For example, regression-in probability was higher for indicative than subjunctive sentences in the spillover region of the comprehension task, but not the judgment task. Considering that several reading measures, including gaze duration and go-past time, indicated target-like sensitivity to the ungrammaticality of the indicative, it is important to interpret these findings with caution since it could be the case that further reanalysis was not required due to rapid detection of earlier integration difficulties, as previously suggested.

In light of the evidence presented above, we argue that task demands and thus reading goals played a central role in modulating processing patterns in polarity contexts. In particular, we found that L2 speakers were more likely to exhibit target-like sensitivity when required to read for judgment than for comprehension. A slight caveat in this respect is that we did not detect any sensitivity to discourse-pragmatic constraints among the L2 learner group. This, however, was in line with our acquisition task predictions (see Section 5.3) formulated within a Feature Reassembly framework (Lardiere, 2009), as part of which we anticipated that L2 learners would treat negated epistemics in a similar manner to subjunctive CP licensing predicates. In other words, that the interlanguage grammar would only license subjunctive CPs as grammatical complements to negated epistemics. However, this is not unexpected when the treatment of the polarity subjunctive in many grammar textbooks aimed at L2 learners of French is considered. For example, most of these textbooks (e.g., Hawkins, Towell and Lamy, 2015) simply state that the subjunctive is the required form following negated epistemics, and never mentions the availability of the indicative in such contexts.

Nevertheless, the current findings (that is, the observation that sensitivity to the ungrammaticality of the indicative under negated epistemics was stronger in the judgment than the comprehension task) are consistent with those presented in Study 3a, despite the focus on different types of subjunctive, and thus substantiate findings from previous studies attesting to more target-like behaviour in tasks that focus on form (Jackson and Bobb, 2009; Leeser, Brandl and Weissglass, 2011; Lim and Christianson, 2013b, 2015). As with Study 3a, we interpret these findings as indicating that L2 learners were more likely to compute the morphologically marked relations in a target-like manner when the task requires a focus on morphosyntactic form. This then suggests that the L1/L2 differences in morphological processing are quantitative rather than qualitative in L2 morphological processing (cf., Jiang, 2004, 2007; Sato and Felser, 2007; Keating, 2009; Lim and Christianson, 2015). For a more in-depth discussion of why this might be the case, see Section 6.5.4.

6.7 Conclusion

To conclude this chapter, the current study has largely shown that the type of secondary task (comprehension questions vs. grammaticality judgment tasks) plays a central role in influencing L2 sensitivity to morphosyntactic violations, regardless of the structure type. In particular, we observed that target-like sensitivity, at least to the ungrammaticality of the indicative but not to the discourse-pragmatic constraints, was more likely to surface in the judgment than the comprehension task. It was argued that these differences in morphological processing were indicative of quantitative rather than qualitative differences in L1 and L2 morphological processing, with reading for judgment leading to deeper processing and more target-like reading behaviour (Lim and Christianson, 2013b, 2015). In this respect, the processing behaviour of the native and non-native speakers was aligned. However, it is important to, particularly in the case of obligatory subjunctives, that it was rarely the secondary task alone that predicted online sensitivity and that factors, such as L1 influence and proficiency, also played a central role. Before concluding, we must recognise an important caveat of these findings relating to the fixed order of the tasks (i.e., the completion of the comprehension task prior to the judgment task). It is clear from the results presented in this chapter that reading times were undoubtedly shorter in the judgment than comprehension task, most likely due to repeated exposure to the experimental stimuli. It is equally possible that this fixed order contributed in large part to the deeper processing patterns observed in the current study. Nevertheless, the current findings are largely consistent previous research showing that target-like sensitivity is more likely to surface in judgment than comprehension tasks (Jackson and Bobb, 2009; Leeser, Brandl and Weissglass, 2011; Lim and Christianson, 2013b, 2015).

Chapter 7 Conclusion

The current thesis has investigated how L2 learners deploy syntactic, semantic and discourse information when processing the subjunctive, using a series of eye-tracking during reading tasks conducted within an anomaly detection paradigm. More specifically, it has considered how L1/L2 differences in featural composition, and its interaction with proficiency, residence abroad and reading goals, influence L2 acquisition and processing.

One of the main contributions of this thesis has been its use of eye-tracking technology. Such a tool has enabled us to develop a more fine-grained understanding of L2 learners' knowledge of the subjunctive. As we have discussed throughout this thesis, eye movement data, more specifically the individual reading measures that index the different levels of processing, has allowed us to explore how specific experimental manipulations affect processing patterns. This is particularly relevant when studying a structure, such as the subjunctive, where successful parsing crucially depends on the ability to integrate syntactic, semantic and discourse information.

Up until this point, we have focused primarily on a series of highly specific hypotheses pertaining to the individual studies presented in each of the chapters that compose the body of this thesis. Let us now return to the overarching research questions outlined in the introduction to this thesis and examine how the data from the individual studies address these questions. Recall the first research question proposed:

RQ1: Do English-speaking L2 learners of French exhibit offline and online sensitivity to the syntactic and semantic constraints of the subjunctive in obligatory contexts? In particular, to what extent is this sensitivity influenced by L1/L2 differences in feature configurations, proficiency and residence in a French-speaking country?

Results from two eye-tracking tasks and a grammaticality judgment task revealed a distinct asymmetry between offline and online knowledge of the subjunctive in obligatory contexts. While offline knowledge of the subjunctive was consistently target-like and in line with previous studies, this was not always the case for online knowledge. We found that online sensitivity was constrained by a multitude of factors including L1 influence, proficiency and residence abroad. In particular, target-like sensitivity was more likely to surface with desire and directive predicates than emotive-factive predicates. We argue that these findings are largely consistent with a processing-based extension of the Feature Reassembly Hypothesis (Lardiere, 2009) and suggest that the acquisition of structures requiring feature re-assembly are more likely to result in non-target-like performance among L2 learners, but that these difficulties can eventually be overcome as proficiency increases. The role of residence abroad, however, was inconclusive, with evidence suggesting that it has both a facilitative and non-facilitative effect on L2 knowledge and processing of the subjunctive.

More generally, these findings have important implications for theories of second language acquisition. As we stated in the introduction to this thesis, some scholars have proposed that L1/L2 differences in featural composition play a deterministic role in L2 learners' sensitivity to morphosyntax during real-time processing (e.g., Dekydtspotter, Schwartz and Sprouse, 2006; Jiang et al., 2011; MacWhinney, 2011; Dekydtspotter and Renaud, 2014), whereas other models do not attribute a significant role to L1 transfer, or at least remain agnostic on the subject (e.g., Clahsen and Felser, 2006, 2018). The current thesis suggests that while theories, such as the Feature Reassembly Hypothesis, have typically been based on findings from a series of offline studies, their predictions equally hold true for online processing patterns. From a methodological perspective, the use of eye-tracking technology has allowed us to gain a more fine-grained insight into the underlying linguistic representations sub-serving second language acquisition, and how these representations can disrupt reading at different levels of processing. Future research within a formal linguistics framework should therefore consider incorporating online methods, such as eye-tracking, more frequently in their experimental design.

Although the current study suggests that the L1 morphosyntax modulates L2 acquisition and processing of the subjunctive, future research is needed to confirm this L1 influence and thus exclude other mitigating factors. For example, it would be advisable to compare L1 English speakers with another L1 group, such as Spanish and Czech, since the former patterns more or less with French in terms of obligatory subjunctive use, whereas the latter does not have a subjunctive mood. Such research would allow us to investigate more directly how L1/L2 morphosyntactic differences influence L2 development and processing.

When discussing the results of Study 1 in Section 4.9, we suggested that limited cognitive resources may in part explain why offline knowledge was not always reflected in online processing patterns. In order to further explore the validity of such a claim, it would be advisable to directly measure individual differences in cognitive resources, such as working memory. Furthermore, we recommend a closer examination of the way in which the input, namely residence abroad, modulates processing patterns given the inconsistent findings presented in this thesis. The current study treated residence abroad as a binary variable and did not appreciate the true complexity that such an experience may have on language development.

RQ2: Do English-speaking L2 learners of French exhibit offline and online sensitivity to the syntactic and discourse/pragmatic constraints of the subjunctive in contexts where the indicative and the subjunctive alternate? To what extent is this knowledge influenced by proficiency and residence in a French-speaking country?

Unlike Study 1, the findings from Study 2 revealed a distinct level of variability among both the L1 and L2 speakers with respect to their knowledge of the polarity subjunctive. In particular, we found an absence of sensitivity to the discourse-pragmatic constraints governing mood use among both experimental groups. This did not appear to be modulated by either proficiency or residence abroad among the L2 learners. We argued that this variability could, in large part, be explained by the tendency to use the subjunctive following negated epistemics, regardless of discourse-pragmatic constraints among L1 speakers, the infrequency of the subjunctive in the input and the cognitive burden that the processing of the polarity subjunctive places on the reader.

It is therefore not surprising, given the distinct variability in the L1 data, that we did not find evidence of target-like sensitivity among the L2 speakers. Again, these findings have important implications for L2 development, especially when considered in conjunction with those of Study 1. In Study 1, we found consistent and most importantly unambiguous evidence of L1 sensitivity to mood-modality mismatches. In other words, the input to which L2 learners were exposed was unambiguous and thus fulfilled Slabakova's (2015) input requirements for target-like acquisition. This was simply not the case with polarity subjunctives; the input for this structure was inconsistent and ambiguous, which appeared to have a subsequent impact on L2 behaviour in the form of indeterminacy between the two moods. Such findings thus reinforce repeated claims in the literature that L2 learners must be exposed to an abundance of unambiguous input, without which successful acquisition is not always guaranteed.

Future research in this area should focus more directly on the potential causes of this variability in the L1 speakers. Understanding why L1 speakers do not always exhibit consistent sensitivity to the discourse-pragmatic constraints particularly during online processing would in turn help us gain a greater understanding of why such a trend should surface among L2 learners as well. Furthermore, as we have suggested, it is likely that the processing of polarity subjunctives places a considerable cognitive burden on the processing systems. It would thus be advisable to examine whether individual differences in the cognitive resources sub-serving the processing systems, such as working memory or pragmatic skill, contribute in any way to the variability observed in this study.

RQ3: To what extent do secondary tasks (i.e., comprehension questions vs. judgment tasks) modulate online sensitivity to the syntactic, semantic and discourse/pragmatic constraints of the subjunctive?

In response to our third and final research question, Study 3 showed that that the type of secondary task (comprehension questions vs. grammaticality judgment tasks) and thus reading goals played a central role in modulating L2 sensitivity to the subjunctive, both in obligatory and polarity contexts. In particular, we found that target-like sensitivity was more likely to surface in the judgment than the comprehension task. We therefore argued that these differences in processing between tasks were indicative of quantitative as opposed to qualitative differences in L1 and L2 morphological processing, with reading for judgment resulting in more deeper processing. However, it must be noted that the type of secondary task alone did not determine L2 sensitivity. In fact, it was often its interaction with other factors, namely the L1 properties and proficiency, that predicted sensitivity. These findings were largely consistent with previous studies showing that target-like sensitivity is more likely to occur in judgment than comprehension tasks (Williams, 2006; Jackson and Bobb, 2009; Leeser, Brandl and Weissglass, 2011; Lim and Christianson, 2013b, 2013a, 2015).

More generally, the findings presented in this thesis have implications for experimental design in L2 research. In particular, the current thesis has shown that the inclusion of secondary tasks, such as comprehension questions and/or judgment tasks, do not have an innocuous effect on reading behaviour. In fact, these tasks can, in large part, determine the detection of sensitivity effects. In the absence of the judgment task, it would have been possible to conclude that L2 sensitivity to mood-modality mismatches during real-time processing was limited at best. We thus encourage researchers to take this finding into account when designing future studies.

However, it is important to recall that the effect of secondary tasks has only been studied in relation to a somewhat limited range of structures. Future research should investigate whether similar effects are present with other grammatical structures. For example, it could be the case, as argued by Godfroid et al. (2015), that knowledge of explicitly taught structures, such as the subjunctive in obligatory contexts, is more likely to surface in judgment tasks due to the focus on form and high levels of metalinguistic knowledge among instructed L2 learners. The fact that target-like sensitivity was more likely to be found in the judgment than the comprehension task with the polarity subjunctive, in a sense, offers a counterexample to this claim, since it is not frequently taught in the L2 classroom. However, further research is required to test the validity of such an argument.

CONCLUDING REMARKS

To summarise, the current thesis has presented findings from several studies investigating the L2 acquisition and processing of the French subjunctive in both obligatory and polarity contexts and the extent to which L1/L2 differences in feature configurations, proficiency, residence in a French-speaking country and reading goals influence both offline and online sensitivity. The key findings can be summarised as follows:

- While L2 learners exhibited offline sensitivity to mood-modality mismatches in obligatory contexts, this was not always the case with online sensitivity. In particular, we found that online sensitivity was highly constrained by multiple factors, including the L1 properties, proficiency and residence abroad.
- There was a distinct absence of sensitivity to the discourse-pragmatic constraints of the subjunctive among both the L1 and L2 speakers. Unlike with obligatory subjunctives, this was not modulated by proficiency and/or residence abroad among the L2 speakers.
- Sensitivity was more likely to surface in tasks that require L2 learners to focus on form rather than comprehension.

Appendix A Test Items

A.1 Study 0: English Subjunctive

CONDITION: DESIRE

Item 1

a. The doctor wants Pierre to be more aware of the risks associated with smoking.
b. The doctor wants that Pierre is more aware of the risks associated with smoking.
c. The doctor wants that Pierre be more aware of the risks associated with smoking.

Item 2

a. Marie wants her daughter to get good grades in her exams.
b. Marie wants that her daughter gets good grades in her exams.
c. Marie wants that her daughter get good grades in her exams.

Item 3

a. Cécilia wants her friend to have more confidence in her (abilities).
b. Cécilia wants that her friend has more confidence in her (abilities).
c. Cécilia wants that her friend have more confidence in her (abilities).

Item 4

a. My housemate prefers the apartment to be tidy.
b. My housemate prefers that the apartment is tidy.
c. My housemate prefers that the apartment be tidy.

Item 5

a. Pierre prefers the meeting to take place at the beginning of the year.
b. Pierre prefers that the meeting takes place at the beginning of the year.
c. Pierre prefers that the meeting take place at the beginning of the year.

Item 6

a. Victoria prefers Claude to be in charge of the project.
b. Victoria prefers that Claude is in charge of the project.
c. Victoria prefers that Claude be in charge of the project.

CONDITION: DIRECTIVE

Item 7

a. The mayor demands Trump to not be allowed to visit the country.
b. The mayor demands that Trump is not allowed to visit the country.
c. The mayor demands that Trump not be allowed to visit the country.

Item 8

a. The minister demands the conference to take place before the end of the year.
b. The minister demands that the conference takes place before the end of the year.
c. The minister demands that the conference take place before the end of the year.

Item 9

a. Switzerland requests the organisation to be equipped with enough resources to fulfil its duties.
b. Switzerland requests that the organisation is equipped with enough resources to fulfil its duties.
c. Switzerland requests that the organisation be equipped with enough resources to fulfil its duties.

Item 10

a. The organisation requests housing to be made available in every neighbourhood.
b. The organisation requests that housing is made available in every neighbourhood.
c. The organisation requests that housing be made available in every neighbourhood.

Item 11

a. The doctor recommends for your diet to be full of fruit and vegetables.
b. The doctor recommends that your diet is full of fruit and vegetables.
c. The doctor recommends that your diet be full of fruit and vegetables.

Item 12

a. The committee recommends for Egypt to make an effort to abolish the death penalty.
b. The committee recommends that Egypt makes an effort to abolish the death penalty.
c. The committee recommends that Egypt make an effort to abolish the death penalty.

CONDITION: EMOTIVE-FACTIVE

Item 13

a. The dentist likes his patient to be aware of the risks associated with eating sugar.
b. The dentist likes that his patient is aware of the risks associated with eating sugar.
c. The dentist likes that his patient be aware of the risks associated with eating sugar.

Item 14

a. The student likes his teacher to make an effort to remember the names of everyone in his class.
b. The student likes that his teacher makes an effort to remember the names of everyone in his class.
c. The student likes that his teacher make an effort to remember the names of everyone in his class.

Item 15

a. Marie regrets Pierre to be late for Mathilde's party.
b. Marie regrets that Pierre is late for Mathilde's party.
c. Marie regrets that Pierre be late for Mathilde's party.

Item 16

a. The minister regrets for the rally to be the target of a political movement.
b. The minister regrets that the rally is the target of a political movement.
c. The minister regrets that the rally be the target of a political movement.

Item 17

a. The government is afraid for the future of the project to be in danger.
b. The government is afraid that the future of the project is in danger.
c. The government is afraid that the future of the project be in danger.

Item 18

a. Pierre is afraid for Sophie to be still in love with Claude.
b. Pierre is afraid that Sophie is still in love with Claude.
c. Pierre is afraid that Sophie be still in love with Claude.

A.2 Study 1: Obligatory Subjunctives

A.2.1 Lexical Properties of Test Items

Table 7.1 Lexical characteristics of test items. Means (standard deviations)

	Desire		Directive		Emotive-Factive	
	Indicative	Subjunctive	Indicative	Subjunctive	Indicative	Subjunctive
Sentence Length (in Characters)	68.08 (12.91)	69.33 (12.72)	76.50 (12.84)	77.83 (12.78)	71.67 (15.47)	72.75 (15.43)
Sentence Length (in Words)	10.75 (2.34)	10.75 (2.34)	12.50 (1.57)	12.50 (1.57)	12.08 (2.97)	12.08 (2.97)
Critical Region (in Characters)	2.58 (1.00)	3.83 (0.58)	2.83 (1.19)	4.17 (0.94)	3.00 (0.74)	4.08 (0.51)
Number of Morphemes	1.00 (0.00)	1.00 (0.00)	1.00 (0.00)	1.00 (0.00)	1.00 (0.00)	1.00 (0.00)
Number of Phonemes	1.08 (0.29)	2.50 (0.90)	1.33 (0.49)	2.50 (0.90)	1.17 (0.39)	2.83 (0.58)
Orthographic Distance to Neighbour	1.31 (0.21)	1.08 (0.05)	1.23 (0.20)	1.15 (0.21)	1.35 (0.18)	1.10 (0.04)
Phonological Distance to Neighbour	1.00 (0.00)	1.00 (0.00)	1.00 (0.00)	1.00 (0.00)	1.00 (0.00)	1.00 (0.00)
Number of Orthographic Neighbours	9.92 (9.38)	14.92 (0.29)	11.92 (8.67)	13.83 (3.13)	7.08 (6.43)	14.83 (0.39)
Number of Phonological Neighbours	18.83 (2.89)	19.50 (2.71)	21.17 (4.71)	21.00 (4.24)	19.67 (3.89)	20.42 (3.55)
Zipf Frequency (Film Subtitles)	7.08 (0.32)	5.31 (0.29)	6.85 (0.42)	5.14 (0.34)	7.10 (0.35)	5.36 (0.25)
Zipf Frequency (Books)	6.60 (0.29)	5.05 (0.25)	6.42 (0.35)	4.92 (0.28)	6.65 (0.28)	5.11 (0.21)

Table 7.2 ANOVA on lexical characteristics of test items across conditions

	Df	Sum Sq	Mean Sq	F value	*p* value
sentence length (in characters)					
condition	1.00	27.00	26.90	0.14	0.71
matrix semantic property	2.00	867.00	433.50	2.29	0.11
condition: matrix semantic property	2.00	0.00	0.10	0.00	1.00
length of critical region in (characters)					
condition	1.00	26.89	26.89	36.34	< 0.01
matrix semantic property	2.00	1.58	0.79	1.07	0.35
condition: matrix semantic property	2.00	0.19	0.10	0.13	0.88
number of morphemes					
condition	1.00	0.00	0.00	1.00	0.32
matrix semantic property	2.00	0.00	0.00	1.00	0.37
condition: matrix semantic property	2.00	0.00	0.00	1.00	0.37
number of phonemes					
condition	1.00	36.12	36.12	88.58	< 0.01
matrix semantic property	2.00	0.53	0.26	0.65	0.53
condition: matrix semantic property	2.00	0.75	0.37	0.92	0.40
orthographic distance to neighbour					
condition	1.00	0.65	0.65	23.60	< 0.01
matrix semantic property	2.00	0.02	0.01	0.37	0.69
condition: matrix semantic property	2.00	0.10	0.05	1.76	0.18
phonological distance to neighbour					
condition	1.00	0.00	0.00	1.00	0.32
matrix semantic property	2.00	0.00	0.00	1.00	0.37
condition: matrix semantic property	2.00	0.00	0.00	1.00	0.37
no. of orthographic neighbours					
condition	1.00	430.20	430.20	12.04	< 0.01
matrix semantic property	2.00	48.10	24.00	0.67	0.51
condition: matrix semantic property	2.00	102.20	51.10	1.43	0.25
no. of phonological neighbours					
condition	1.00	3.10	3.13	0.22	0.64
matrix semantic property	2.00	44.20	22.10	1.59	0.21
condition: matrix semantic property	2.00	3.10	1.54	0.11	0.90
zipf frequency (film subtitles)					
condition	1.00	54.43	54.43	489.07	< 0.01
matrix semantic property	2.00	0.81	0.40	3.63	0.03
condition: matrix semantic property	2.00	0.01	0.01	0.06	0.94
zipf frequency (books)					
condition	1.00	42.13	42.13	534.23	< 0.01
matrix semantic property	2.00	0.56	0.28	3.57	0.03
condition: matrix semantic property	2.00	0.01	0.00	0.06	0.94

Note: Shaded cells denote statistical significance ($p < 0.05$).

A.2.2 Mood Morphology Production Task

Conjuguez les verbes au mode qui convient.

1. Les enfants sont contents qu'Eurodisney __________ (être) ouvert
2. Marie est déçue qu'il ne __________ (faire) pas beau aujourd'hui
3. Paul est malheureux que ses amies __________ (devoir) partir
4. Elle est soulagée que ses cousins __________ (avoir) réussi
5. Il vaut mieux que vous __________ (rester) avec votre cousine
6. Il était temps que vous __________ (arriver) enfin
7. Il sera essentiel que vous __________ (répondre) à toutes les questions
8. Il était juste que tous __________ (avoir) reçu la même note
9. Il faudra que l'article __________ (paraître)demain matin
10. Il est essentiel que tous les documents __________ (être) signés
11. Mes parents ont interdit que nous __________ (sortir) le soir
12. Paul a défendu que vous __________ (parler) sans un avocat
13. Marie ne permettra pas que les enfants __________ (partir) seuls
14. J'aurais voulu que vous y __________ (aller) avec moi
15. Nous souhaitons tous que le mariage __________ (avoir) lieu en juin
16. Désires-tu que les fleurs __________ (être) livrées ce matin?
17. Je doute que vous __________ (savoir) la vérité, c'est un secret
18. Il est possible que nous __________ (perdre) le match
19. Il était impossible que vous __________ (gagner) ce set contre Pierre
20. Il se peut que Marie n'__________ (avoir) pas lu ta lettre
21. Il n'est pas sûr que les étudiants se __________ (être) tous inscrits hier.
22. Il est improbable que nous __________ (voyager) l'été prochain.

A.2.3 Lextale-Fr

Stimulus	Mot?	Stimulus	Mot?	Stimulus	Mot?
cheveux		gloque		bouton	
soumon		lézard		capeline	
cloche		sacher		lanière	
fascine		nouer		honteur	
huif		occire		abêtir	
semonce		écouce		fenêtre	
canoter		osseaux		écureuil	
infâme		rejoute		caddie	
fourmi		escroc		détume	
cadenas		hache		oeuiller	
racaille		parchance		balai	
pourcine		pinceau		prioche	
œillet		poisson		vicelard	
raplaner		robinet		joueux	
plaiser		amadouer		agire	
cerveler		peigne		éventail	
endifier		retruire		boutard	
jamain		crayon		panier	
ennemi		sentuelle		citrouille	
pouce		alourdir		bouilloire	
metter		marteau		parir	
fosse		esquif		remporter	
inciter		treillage		procoreux	
salière		dauphin		tanin	
fouet		orgueil		église	
cessure		amorce		indicible	
clouer		cintre		réporce	
mappemonde		chameau		mignon	

A.2.4 Language Background Questionnaire

Language background questionnaire for final year students

Section 1. Background Information

Question 1.1

Name

Question 1.2

Date of Birth:

1 / January / 2018

Question 1.3

Sex:

- Male
- Female
- Other

Question 1.4

Education (your current or most recent education level, even if you have not finished the degree)

- Postgraduate (PhD/MD/JD)
- Postgraduate (Masters)
- Undergraduate (BA/BSc)
- A-Levels
- GCSEs
- Other

Question 1.4b

If you selected "other", please specify here:

Question 1.5

Degree Title

Question 1.6

Are you currently enrolled in a four-year undergraduate degree programme with a compulsory year abroad in the Modern Languages and Linguistics department at the University of Southampton?

- Yes

- No

Question 1.6b

What languages do you study? Please also indicate the stage too.

Question 1.6c

Indicate your year of study

- First year
- Second year
- Third year
- Fourth year

Question 1.7

Handedness

- Left
- Right

Question 1.8

City and country of residence:

Question 1.9

City and country of origin:

Question 1.10

If 6 and 7 are different, then when did you first move to the country where you currently live?

Question 1.11

Do you have any vision problems?

- Yes
- No

Question 1.11b

Do you wear glasses or contacts?

- Yes
- No

Question 1.12

Indicate your native language(s) and any other languages you have studied or learned, the age at which you started using each language in terms of listening, speaking, reading, and writing, and the total number of years you have spent using each language.

If you speak more than one language, separate each response with //

N.B. You may have learned a language, stopped using it, and then started using it again. Please give the total number of years.

Language
Listening
Speaking
Reading
Writing
Years of use

Question 1.13

Have you ever studied or learned a second language in terms of listening, speaking, reading, or writing?

- Yes
- No

Question 1.13b

Indicate the age at which you started using each of the languages you have studied or learned in the following environments.

If you have studied or learned more than one language, separate each response with //

Language
At home
With friends
At school
At work
Language software
Online

Question 1.13c

Rate on a scale of 1-7 your current ability in terms of listening, speaking, reading, and writing in each of the languages you have studied or learned.

(1 = very poor, 2 = poor, 3 = limited, 4 = functional, 5 = good, 6 = very good, 7 = native-like)

If you have studied or learned more than one language, separate each response with //

Language
Listening
Speaking
Reading
Writing

Question 1.13d

Estimate how many hours per day you spend engaged in the following activities in each of the languages you have studied or learned.

If you have studied or learned more than one language, separate each response with //

Watching television
Listening to radio

Reading for fun
Reading for school/work
Writing emails to friends
Writing for school/work

Question 1.13e

Estimate how many hours per day you spend speaking with the following groups of people in each of the languages you have studied or learned.

If you have studied or learned more than one language, separate each response with //

Include anyone in the work environment in the 'coworkers' category (e.g., if you are a teacher

Family members
Significant others
Friends
Classmates
Coworkers

Question 1.13f

How often do you use each of the languages you have studied or learned for the following activities?

If you speak more than one language, separate each response with //

(1 = never, 2 = rarely, 3 = sometimes, 4 = regularly, 5 = often, 6 = usually, 7 = always)

Language
Thinking
Talking to yourself
Expression emotion (e.g. shouting cursing showing affection etc.)
Dreaming
Arithmetic (e.g. counting calculating tips etc.)
Remember numbers (e.g. telephone numbers ID numbers etc.)

Question 1.13g

What percentage of your friends speaks each of the languages you have studied or learned? *(The total percentage should add up to 100%.)*

If you have studied or learned more than one language, separate each response with //

Language
Percentage

Question 1.14

Have you lived or travelled in countries other than your country of residence or country of origin for three or more months?

(If you select yes, you will be able to enter up to 4 countries)

- ◯ Yes
- ◯ No

Question 1.14b

Indicate the name of the country

Question 1.14c

Indicative the length of stay [month(s)]

N.B. You may have been to the country on multiple occasions, each for a different length of time. Add all the trips together.

Question 1.14d

Indicate the language you used

Question 1.14e

Indicate the frequency of use of the language

(1 = never, 2 = rarely, 3 = sometimes, 4 = regularly, 5 = often, 6 = usually, 7 = always)

Never

- ◯ 1
- ◯ 2
- ◯ 3
- ◯ 4
- ◯ 5
- ◯ 6
- ◯ 7

Always

Question 1.14f

Do you wish to add another country?

- ◯ Yes
- ◯ No

Section 2. Final Year Students

Question 2.1

What language did you speak during your year abroad?

Question 2.2

How often did you do the following during your year abroad in the dominant language?

	Every day	Several times a week	A few times a week	A couple of times a month	Rarely	Never

Watch TV	○	○	○	○	○	○
Watch films	○	○	○	○	○	○
Browse the internet (e.g. read the news, etc.)	○	○	○	○	○	○
Use social networking sites (e.g. Facebook/Twitter)	○	○	○	○	○	○
Read emails	○	○	○	○	○	○
Write emails	○	○	○	○	○	○
Listen to music	○	○	○	○	○	○
Watch TV	○	○	○	○	○	○
Watch films	○	○	○	○	○	○
Browse the internet (eg. read news, etc)	○	○	○	○	○	○
Use social networking sites (eg. Facebook/ Twitter)	○	○	○	○	○	○
Read emails	○	○	○	○	○	○
Write emails	○	○	○	○	○	○
Listen to music	○	○	○	○	○	○
Listen to talk radio	○	○	○	○	○	○
Listen to lectures	○	○	○	○	○	○
Participate in seminars/ language classes	○	○	○	○	○	○
Read literature (eg. fiction, poetry, short stories)	○	○	○	○	○	○
Read academic texts	○	○	○	○	○	○
Read newspapers	○	○	○	○	○	○
Read magazines	○	○	○	○	○	○
Read text messages	○	○	○	○	○	○
Write text messages	○	○	○	○	○	○
Write reports (eg. work, academic)	○	○	○	○	○	○
Write for leisure (eg. journal)	○	○	○	○	○	○
Use instant messaging	○	○	○	○	○	○
Have short phone/ Skype/ etc conversations (<5 minutes)	○	○	○	○	○	○
Have long phone/ Skype/ etc conversations (>5 minutes)	○	○	○	○	○	○
Teach a class	○	○	○	○	○	○
Engage in service encounters	○	○	○	○	○	○
Engage in small talk	○	○	○	○	○	○
Engage in long casual conversations	○	○	○	○	○	○
Participate in organised social activities (eg, clubs, church, sports, etc)	○	○	○	○	○	○
Have meetings	○	○	○	○	○	○

Thank you for taking this questionnaire.

A.2.5 Experimental Items

DESIRE

Item 1

a. Le docteur veut que Pierre soit plus conscient des risques associés à la consommation du tabac.
b. Le docteur veut que Pierre est plus conscient des risques associés à la consommation du tabac.

Item 2

a. Marie veut que son enfant a une mention très bien au baccalauréat.
b. Marie veut que son enfant ait une mention très bien au baccalauréat.

Item 3

a. Cécilia veut que son amie ait plus de confiance en elle.
b. Cécilia veut que son amie a plus de confiance en elle.

Item 4

a. Ma colocataire préfère que l'appartement est toujours propre.
b. Ma colocataire préfère que l'appartement soit toujours propre.

Item 5

a. Pierre préfère que la conférence a lieu en début d'année.
b. Pierre préfère que la conférence ait lieu en début d'année.

Item 6

a. Victoria préfère que Claude soit responsable du projet.
b. Victoria préfère que Claude est responsable du projet.

Item 7

a. Manon préfère que Thibault fait des études supérieures.
b. Manon préfère que Thibault fasse des études supérieures.

Item 8

a. Mathieu souhaite que Corbyn est candidat aux élections législatives.
b. Mathieu souhaite que Corbyn soit candidat aux élections législatives.

Item 9

a. Le ministre souhaite que le bilinguisme soit perçu comme une richesse.
b. Le ministre souhaite que le bilinguisme est perçu comme une richesse.

Item 10

a. La région souhaite que l'apprentissage est la voie normale d'accès à l'emploi
b. La région souhaite que l'apprentissage soit la voie normale d'accès à l'emploi

Item 11

a. Le directeur souhaite que l'élève soit attentif lors des apprentissages en classe
b. Le directeur souhaite que l'élève est attentif lors des apprentissages en classe

Item 12

a. Macron souhaite que la France soit l'un des leaders de l'intelligence artificielle.
b. Macron souhaite que la France est l'un des leaders de l'intelligence artificielle.

DIRECTIVE

Item 13

a. Le maire exige que Trump soit interdit de visite officielle.
b. Le maire exige que Trump est interdit de visite officielle.

Item 14

a. Le président exige que sa femme a l'image d'une femme au foyer américaine modèle
b. Le président exige que sa femme ait l'image d'une femme au foyer américaine modèle

Item 15

a. Le critère exige que la personne ait une compréhension élémentaire d'une langue étrangère.
b. Le critère exige que la personne a une compréhension élémentaire d'une langue étrangère.

Item 16

a. Le ministre exige que le référendum a lieu avant la fin de l'année.
b. Le ministre exige que le référendum ait lieu avant la fin de l'année.

Item 17

a. Le curriculum exige que chaque élève fasse 40 heures de bénévolat dans sa communauté.
b. Le curriculum exige que chaque élève fait 40 heures de bénévolat dans sa communauté.

Item 18

a. Trump demande que les drapeaux sont en berne pour John McCain.
b. Trump demande que les drapeaux soient en berne pour John McCain.

Item 19

a. La Suisse demande que l'organisation soit dotée de suffisamment de ressources pour accomplir ses mandats.
b. La Suisse demande que l'organisation est dotée de suffisamment de ressources pour accomplir ses mandats.

Item 20

a. Amnesty demande que la police fait preuve de modération lors des manifestations.
b. Amnesty demande que la police fasse preuve de modération lors des manifestations.

Item 21

a. L'organisation recommande que le logement est disponible dans tous les quartiers.
b. L'organisation recommande que le logement soit disponible dans tous les quartiers.

Item 22

a. La commission recommande que le ministre soit déchu de son mandat.
b. La commission recommande que le ministre est déchu de son mandat.

Item 23

a. Le docteur recommande que son régime est riche en fruits et légumes.
b. Le docteur recommande que son régime soit riche en fruits et légumes.

Item 24

a. Le Comité recommande que l'Égypte fasse des efforts pour abolir la peine de mort.
b. Le Comité recommande que l'Égypte fait des efforts pour abolir la peine de mort.

EMOTIVE-FACTIVE

Item 25

a. Le dentiste aime que son patient soit conscient des risques associés à la consommation du sucre.
b. Le dentiste aime que son patient est conscient des risques associés à la consommation du sucre.

Item 26

a. Le lycéen aime que son professeur fait un effort pour se souvenir des noms de toute sa classe.
b. Le lycéen aime que son professeur fasse un effort pour se souvenir des noms de toute sa classe.

Item 27

a. Mélenchon regrette que Macron est le porte-parole de la droite.
b. Mélenchon regrette que Macron soit le porte-parole de la droite.

Item 28

a. Jean-Marc regrette que la jeunesse soit moins politisée qu'au cours des années 1980.
b. Jean-Marc regrette que la jeunesse est moins politisée qu'au cours des années 1980.

Item 29

a. Marie regrette que Pierre est en retard pour la soirée de Mathilde.
b. Marie regrette que Pierre soit en retard pour la soirée de Mathilde.

Item 30

a. Le maire regrette que la mobilisation fasse l'objet d'un mouvement politique.
b. Le maire regrette que la mobilisation fait l'objet d'un mouvement politique.

Item 31

a. Quentin est content que Marie a du temps pour partir en vacances.

b. Quentin est content que Marie ait du temps pour partir en vacances.

Item 32

a. Pierre est content que Nicolas soit en vacances pendant deux semaines.

b. Pierre est content que Nicolas est en vacances pendant deux semaines.

Item 33

a. Lucie est triste que l'été est presque fini.

b. Lucie est triste que l'été soit presque fini.

Item 34

a. Le gouvernement craint que l'avenir du projet soit en danger.

b. Le gouvernement craint que l'avenir du projet est en danger.

Item 35

a. Pierre craint que Sophie est toujours amoureuse de Claude.

b. Pierre craint que Sophie soit toujours amoureuse de Claude.

Item 36

a. Le président craint que son club soit un enjeu dans le cadre des élections municipales.

b. Le président craint que son club est un enjeu dans le cadre des élections municipales.

A.3 Study 2: Polarity Subjunctives

A.3.1 Lexical Properties of Test Items

Table 7.3 Lexical characteristics of test items. Means (standard deviations)

condition	Indicative-False	Subjunctive-False	Indicative-True	Subjunctive-True
sentence length (in characters)	79.31 (5.55)	80.63 (5.56)	73.31 (5.55)	74.61 (5.56)
sentence length (in words)	17.31 (1.05)	17.31 (1.05)	15.31 (1.05)	15.31 (1.05)
critical region (in characters)	2.59 (1.23)	3.94 (0.79)	2.59 (1.23)	3.94 (0.79)
no. of morphemes	1.00 (0.00)	1.00 (0.00)	1.00 (0.00)	1.00 (0.00)
no. of phonemes	1.28 (0.45)	2.31 (0.95)	1.28 (0.45)	2.31 (0.95)
orthographic distance to neighbour	1.20 (0.20)	1.08 (0.06)	1.20 (0.20)	1.08 (0.06)
phonological distance to neighbour	1.00 (0.00)	1.00 (0.00)	1.00 (0.00)	1.00 (0.00)
number of orthographic neighbours	13.47 (8.99)	14.72 (0.45)	13.47 (8.99)	14.72 (0.45)
number of phonological neighbours	20.81 (4.52)	21.19 (4.30)	20.81 (4.52)	21.19 (4.30)
zipf frequency (film subtitles)	6.86 (0.35)	5.14 (0.28)	6.86 (0.35)	5.14 (0.28)
zipf frequency (film subtitles)	6.42 (0.30)	4.92 (0.24)	6.42 (0.30)	4.92 (0.24)

Table 7.4 ANOVA comparing conditions

	Df	Sum Sq	Mean Sq	F value	*p* value
Sentence Length (in Characters)					
Mood	1.00	164.00	164.00	5.32	0.02
Truth	1.00	3462.00	3462.00	112.19	< 0.01
Mood: Truth	1.00	0.00	0.00	0.00	0.99
Length of Critical Region (in Characters)					
Mood	1.00	173.30	173.34	162.40	< 0.01
Truth	1.00	0.00	0.00	0.00	1.00
Mood: Truth	1.00	0.00	0.00	0.00	1.00
Number of Morphemes					
Mood	1.00	0.00	0.00	1.00	0.32
Truth	1.00	0.00	0.00	1.00	0.32
Mood: Truth	1.00	0.00	0.00	1.00	0.32
Number of Phonemes					
Mood	1.00	102.10	102.09	182.90	< 0.01
Truth	1.00	0.00	0.00	0.00	1.00
Mood: Truth	1.00	0.00	0.00	0.00	1.00
Orthographic Distance to Neighbour					
Mood	1.00	1.32	1.32	59.68	0.00
Truth	1.00	0.00	0.00	0.00	1.00
Mood: Truth	1.00	0.00	0.00	0.00	1.00
Phonological Distance to Neighbour					
Mood	1.00	0.00	0.00	1.00	0.32
Truth	1.00	0.00	0.00	1.00	0.32
Mood: Truth	1.00	0.00	0.00	1.00	0.32
No. of Orthographic Neighbours					
Mood	1.00	150.00	150.00	3.71	0.06
Truth	1.00	0.00	0.00	0.00	1.00
Mood: Truth	1.00	0.00	0.00	0.00	1.00
No. of Phonological Neighbours					
Mood	1.00	13.00	13.50	0.69	0.41
Truth	1.00	0.00	0.00	0.00	1.00
Mood: Truth	1.00	0.00	0.00	0.00	1.00
Zipf Frequency (Film Subtitles)					
Mood	1.00	282.91	282.90	2769.00	< 0.01
Truth	1.00	0.00	0.00	0.00	1.00
Mood: Truth	1.00	0.00	0.00	0.00	1.00
Zipf Frequency (Books)					
Mood	1.00	217.86	217.86	2996.00	< 0.01
Truth	1.00	0.00	0.00	0.00	1.00
Mood: Truth	1.00	0.00	0.00	0.00	1.00

Note: Shaded cells denote statistical significance ($p < 0.05$).

A.3.2 Experimental Items

Item 1

a. Marie ne pense pas que Pierre est content et je ne le pense pas non plus.
b. Marie ne pense pas que Pierre est content mais moi je pense que si.
c. Marie ne pense pas que Pierre soit content et je ne le pense pas non plus.
d. Marie ne pense pas que Pierre soit content mais moi je pense que si.

Item 2

a. Manon ne pense pas que Nicholas soit triste mais moi je pense que si.
b. Manon ne pense pas que Nicholas est triste et je ne le pense pas non plus.
c. Manon ne pense pas que Nicholas est triste mais moi je pense que si.
d. Manon ne pense pas que Nicholas soit triste et je ne le pense pas non plus.

Item 3

a. Maxime ne pense pas que Marine fasse la vaisselle et je ne le pense pas non plus.
b. Maxime ne pense pas que Marine fasse la vaisselle mais moi je pense que si.
c. Maxime ne pense pas que Marine fait la vaisselle et je ne le pense pas non plus.
d. Maxime ne pense pas que Marine fait la vaisselle mais moi je pense que si.

Item 4

a. Quentin ne pense pas que Julien est en colère mais moi je pense que si.
b. Quentin ne pense pas que Julien soit en colère et je ne le pense pas non plus.
c. Quentin ne pense pas que Julien soit en colère mais moi je pense que si.
d. Quentin ne pense pas que Julien est en colère et je ne le pense pas non plus.

Item 5

a. Léa ne pense pas que Guillaume fait du vélo et je ne le pense pas non plus.
b. Léa ne pense pas que Guillaume fait du vélo mais moi je pense que si.
c. Léa ne pense pas que Guillaume fasse du vélo et je ne le pense pas non plus.
d. Léa ne pense pas que Guillaume fasse du vélo mais moi je pense que si.

Item 6

a. Mathilde ne pense pas que Valentin ait beaucoup d'argent mais moi je pense que si.
b. Mathilde ne pense pas que Valentin a beaucoup d'argent et je ne le pense pas non plus.
c. Mathilde ne pense pas que Valentin a beaucoup d'argent mais moi je pense que si.
d. Mathilde ne pense pas que Valentin ait beaucoup d'argent et je ne le pense pas non plus.

Item 7

a. Mathieu ne pense pas que Morgane ait les cheveux bruns et je ne le pense pas non plus.
b. Mathieu ne pense pas que Morgane ait les cheveux bruns mais moi je pense que si.
c. Mathieu ne pense pas que Morgane a les cheveux bruns et je ne le pense pas non plus.
d. Mathieu ne pense pas que Morgane a les cheveux bruns mais moi je pense que si.

Item 8

a. Sébastien ne pense pas que Clémence a les cheveux blonds mais moi je pense que si.
b. Sébastien ne pense pas que Clémence ait les cheveux blonds et je ne le pense pas non plus.
c. Sébastien ne pense pas que Clémence ait les cheveux blonds mais moi je pense que si.
d. Sébastien ne pense pas que Clémence a les cheveux blonds et je ne le pense pas non plus.

Item 9

a. Juliette ne pense pas que le chat a les yeux verts et je ne le pense pas non plus.
b. Juliette ne pense pas que le chat a les yeux verts mais moi je pense que si.
c. Juliette ne pense pas que le chat ait les yeux verts et je ne le pense pas non plus.
d. Juliette ne pense pas que le chat ait les yeux verts mais moi je pense que si.

Item 10

a. Maeva ne pense pas que le chat ait les yeux bleus mais moi je pense que si.
b. Maeva ne pense pas que le chat a les yeux bleus et je ne le pense pas non plus.
c. Maeva ne pense pas que le chat a les yeux bleus mais moi je pense que si.
d. Maeva ne pense pas que le chat ait les yeux bleus et je ne le pense pas non plus.

Item 11

a. François ne pense pas que Cécilia fasse du shopping et je ne le pense pas non plus.
b. François ne pense pas que Cécilia fasse du shopping mais moi je pense que si.
c. François ne pense pas que Cécilia fait du shopping et je ne le pense pas non plus.
d. François ne pense pas que Cécilia fait du shopping mais moi je pense que si.

Item 12

a. Laurent ne pense pas que Marie est en vacances mais moi je pense que si.
b. Laurent ne pense pas que Marie soit en vacances et je ne le pense pas non plus.
c. Laurent ne pense pas que Marie soit en vacances mais moi je pense que si.
d. Laurent ne pense pas que Marie est en vacances et je ne le pense pas non plus.

Item 13

a. Émilie ne pense pas que Tom fait de la natation et je ne le pense pas non plus.
b. Émilie ne pense pas que Tom fait de la natation mais moi je pense que si.
c. Émilie ne pense pas que Tom fasse de la natation et je ne le pense pas non plus.
d. Émilie ne pense pas que Tom fasse de la natation mais moi je pense que si.

Item 14

a. Margot ne pense pas que Felix soit à Paris mais moi je pense que si.
b. Margot ne pense pas que Felix est à Paris et je ne le pense pas non plus.
c. Margot ne pense pas que Felix est à Paris mais moi je pense que si.
d. Margot ne pense pas que Felix soit à Paris et je ne le pense pas non plus.

Item 15

a. Capucine ne pense pas que Robin soit au cinéma et je ne le pense pas non plus.
b. Capucine ne pense pas que Robin soit au cinéma mais moi je pense que si.
c. Capucine ne pense pas que Robin est au cinéma et je ne le pense pas non plus.
d. Capucine ne pense pas que Robin est au cinéma mais moi je pense que si.

Item 16

a. Maëlle ne pense pas que Florent est à l'aéroport mais moi je pense que si.
b. Maëlle ne pense pas que Florent soit à l'aéroport et je ne le pense pas non plus.
c. Maëlle ne pense pas que Florent soit à l'aéroport mais moi je pense que si.
d. Maëlle ne pense pas que Florent est à l'aéroport et je ne le pense pas non plus.

Item 17

a. Amandine ne croit pas que Christophe fait du snowboard et je ne le crois pas non plus.
b. Amandine ne croit pas que Christophe fait du snowboard mais moi je crois que si.
c. Amandine ne croit pas que Christophe fasse du snowboard et je ne le crois pas non plus.
d. Amandine ne croit pas que Christophe fasse du snowboard mais moi je crois que si.

Item 18

a. Elodie ne croit pas que Ruben soit à la plage mais moi je crois que si.
b. Elodie ne croit pas que Ruben est à la plage et je ne le crois pas non plus.
c. Elodie ne croit pas que Ruben est à la plage mais moi je crois que si.
d. Elodie ne croit pas que Ruben soit à la plage et je ne le crois pas non plus.

Item 19

a. Sophie ne croit pas que Jean soit pompier et je ne le crois pas non plus.
b. Sophie ne croit pas que Jean soit pompier mais moi je crois que si.
c. Sophie ne croit pas que Jean est pompier et je ne le crois pas non plus.
d. Sophie ne croit pas que Jean est pompier mais moi je crois que si.

Item 20

a. Margaux ne croit pas que Leo est médecin mais moi je crois que si.
b. Margaux ne croit pas que Leo soit médecin et je ne le crois pas non plus.
c. Margaux ne croit pas que Leo soit médecin mais moi je crois que si.
d. Margaux ne croit pas que Leo est médecin et je ne le crois pas non plus.

Item 21

a. Maxime ne croit pas que Juliette est endormie et je ne le crois pas non plus.
b. Maxime ne croit pas que Juliette est endormie mais moi je crois que si.
c. Maxime ne croit pas que Juliette soit endormie et je ne le crois pas non plus.
d. Maxime ne croit pas que Juliette soit endormie mais moi je crois que si.

Item 22

a. Alexandre ne croit pas que Jeanne soit au travail mais moi je crois que si.
b. Alexandre ne croit pas que Jeanne est au travail et je ne le crois pas non plus.
c. Alexandre ne croit pas que Jeanne est au travail mais moi je crois que si.
d. Alexandre ne croit pas que Jeanne soit au travail et je ne le crois pas non plus.

Item 23

a. Mathilde ne croit pas que Gabriel ait le cœur brisé et je ne le crois pas non plus.
b. Mathilde ne croit pas que Gabriel ait le cœur brisé mais moi je crois que si.
c. Mathilde ne croit pas que Gabriel a le cœur brisé et je ne le crois pas non plus.
d. Mathilde ne croit pas que Gabriel a le cœur brisé mais moi je crois que si.

Item 24

a. Lucie ne croit pas que Nathan a mal à la tête mais moi je crois que si.
b. Lucie ne croit pas que Nathan ait mal à la tête et je ne le crois pas non plus.
c. Lucie ne croit pas que Nathan ait mal à la tête mais moi je crois que si.
d. Lucie ne croit pas que Nathan a mal à la tête et je ne le crois pas non plus.

Item 25

a. Claude ne croit pas que Charlotte fait du foot et je ne le crois pas non plus.
b. Claude ne croit pas que Charlotte fait du foot mais moi je crois que si.
c. Claude ne croit pas que Charlotte fasse du foot et je ne le crois pas non plus.

d. Claude ne croit pas que Charlotte fasse du foot mais moi je crois que si.

Item 26

a. Clément ne croit pas que Mathilde fasse le ménage mais moi je crois que si.
b. Clément ne croit pas que Mathilde fait le ménage et je ne le crois pas non plus.
c. Clément ne croit pas que Mathilde fait le ménage mais moi je crois que si.
d. Clément ne croit pas que Mathilde fasse le ménage et je ne le crois pas non plus.

Item 27

a. Mathis ne croit pas que Sarah ait mal au dos et je ne le crois pas non plus.
b. Mathis ne croit pas que Sarah ait mal au dos mais moi je crois que si.
c. Mathis ne croit pas que Sarah a mal au dos et je ne le crois pas non plus.
d. Mathis ne croit pas que Sarah a mal au dos mais moi je crois que si.

Item 28

a. Lola ne croit pas que Tom a chaud mais moi je crois que si.
b. Lola ne croit pas que Tom ait chaud et je ne le crois pas non plus.
c. Lola ne croit pas que Tom ait chaud mais moi je crois que si.
d. Lola ne croit pas que Tom a chaud et je ne le crois pas non plus.

Item 29

a. Eva ne croit pas que David a froid et je ne le crois pas non plus.
b. Eva ne croit pas que David a froid mais moi je crois que si.
c. Eva ne croit pas que David ait froid et je ne le crois pas non plus.
d. Eva ne croit pas que David ait froid mais moi je crois que si.

Item 30

a. Noémie ne croit pas que Thibault fasse ses devoirs mais moi je crois que si.
b. Noémie ne croit pas que Thibault fait ses devoirs et je ne le crois pas non plus.
c. Noémie ne croit pas que Thibault fait ses devoirs mais moi je crois que si.
d. Noémie ne croit pas que Thibault fasse ses devoirs et je ne le crois pas non plus.

Item 31

a. Chloé ne croit pas que Baptiste soit en retard au bureau et je ne le crois pas non plus.
b. Chloé ne croit pas que Baptiste soit en retard au bureau mais moi je crois que si.
c. Chloé ne croit pas que Baptiste est en retard au bureau et je ne le crois pas non plus.
d. Chloé ne croit pas que Baptiste est en retard au bureau mais moi je crois que si.

Item 32

a. Oscar ne croit pas que Claire fait la bise à son copain mais moi je crois que si.
b. Oscar ne croit pas que Claire fasse la bise à son copain et je ne le crois pas non plus.
c. Oscar ne croit pas que Claire fasse la bise à son copain mais moi je crois que si.
d. Oscar ne croit pas que Claire fait la bise à son copain et je ne le crois pas non plus.

Bibliography

Abouda, L. (2002) 'Négation, Interrogation et Alternance Indicatif–Subjonctif', *Journal of French Language Studies*, 12(1), pp. 1–22. doi: 10.1017/S095926950200011X.

Altarriba, J., Kroll, J. F., Sholl, A. and Rayner, K. (1996) 'The influence of lexical and conceptual constraints on reading mixed language sentences: Evidence from eye fixation and naming times', *Memory & Cognition*, 24, pp. 477–492.

Arnett, C. (2013) 'Syntactic gains in short-term study abroad', *Foreign Language Annals*, 46(4), pp. 705–712. doi: 10.1111/flan.12052.

Ayoun, D. (2013) *The Second Language Acquisition of French Tense, Aspect, Mood and Modality*. Amsterdam: John Benjamins Publishing Company (AILA Applied Linguistics Series). doi: 10.1075/aals.10.

Ayres-Bennett, W. and Carruthers, J. (2001) 'Mood and modality: The French subjunctive', in Ayres-Bennett, W., Carruthers, J., and Temple, R. (eds) *Problems and Perspectives: Studies in the Modern French Language*. London: Routledge, pp. 192–221.

Balota, D. A., Pollatsek, A. and Rayner, K. (1985) 'The interaction of contextual constraints and parafoveal visual information in reading', *Cognitive Psychology*, 17, pp. 364–388.

Barr, D. J., Levy, R., Scheepers, C. and Tily, H. J. (2013) 'Random effects structure for confirmatory hypothesis testing: Keep it maximal', *Journal of Memory and Language*, 68(3), pp. 255–278. doi: 10.1016/j.jml.2012.11.001.

Bartning, I. (2008) 'The Advanced Learner Variety: 10 Years Later', in Labeau, E. and Myles, F. (eds) *Contemporary Studies in Descriptive Linguistics, Volume 12: Advanced Learner Variety: The Case of French*. Oxford: Peter Lang AG, pp. 11–40.

Bartning, I., Lundell, F. F. and Hancock, V. (2012) 'On the Role of Linguistic Contextual Factors for Morphosyntactic Stabilization in High-Level L2 French', *Studies in Second Language Acquisition*, 34(02), pp. 243–267. doi: 10.1017/S0272263112000046.

Bates, D., Mächler, M., Bolker, B. and Walker, S. (2015) 'Fitting Linear Mixed-Effects Models Using lme4', *Journal of Statistical Software*, 67(1). doi: 10.18637/jss.v067.i01.

Baum, S. and Titone, D. A. (2014) 'Moving toward a neuroplasticity view of bilingualism, executive control, and aging', *Applied Psycholinguistics*, 35(5), pp. 857–894. doi: 10.1017/S0142716414000174.

Baunaz, L. (2017) 'Embedding verbs and subjunctive mood: The emotive factor', in Perpiñán, S., Heap, D., and Moreno-Villamar, I. (eds) *Romance Languages and Linguistic Theory 11. Selected papers from the 44th Linguistic Symposium on romance Languages*. Amsterdam: John Benjamins Publishing Company, pp. 9–31. doi: 10.1075/rllt.11.01bau.

Belletti, A., Bennati, E. and Sorace, A. (2007) 'Theoretical and developmental issues in the syntax of subjects: Evidence from near-native Italian', *Natural Language & Linguistic Theory*, 25(4), pp. 657–689. doi: 10.1007/s11049-007-9026-9.

Bialystok, E. (1979) 'Explicit and implicit judgements of L2 grammaticality', *Language Learning*, 29(1), pp. 81–103. doi: 10.1111/j.1467-1770.1979.tb01053.x.

Bialystok, E. (1982) 'On the relationship between knowing and using linguistic forms', *Applied Linguistics*, III(3), pp. 181–206. doi: 10.1093/applin/III.3.181.

Bialystok, E. (1994) 'Analysis and Control in the Development of Second Language Proficiency', *Studies in Second Language Acquisition*, 16(2), pp. 157–168. doi: 10.1017/S0272263100012857.

Bianchi, V. (2001) *On Person agreement.* Pisa.

Black, J., Williams, D. and Ferguson, H. J. (2018) 'Imagining counterfactual worlds in autism spectrum disorder.', *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 44(9), pp. 1444–1463. doi: 10.1037/xlm0000500.

Bley-Vroman, R. (1983) 'The comparative fallacy in interlanguage studies: the case of systematicity.', *Language Learning*, 33(1), pp. 1–17. doi: 10.1111/j.1467-1770.1983.tb00983.x.

Borgonovo, C., Garavito, J. B. De and Prévost, P. (2015) 'Mood selection in relative clauses', *Studies in Second Language Acquisition*, 37(1), pp. 33–69. doi: 10.1017/S0272263114000321.

Borgonovo, C. and Prévost, P. (2003) 'Knowledge of polarity subjunctive in L2 Spanish', in Beachley, B., Brown, A., & Conlin, F. (ed.) *Proceedings of the 27th BUCLD*. Somerville, MA: Cascadilla Press, pp. 150–161.

Bot, K. de (2015) 'Moving Where? A Reaction to Slabakova et al. (2014)', *Applied Linguistics*, 36(2), pp. 261–264. doi: 10.1093/applin/amu074.

Boxell, O. and Felser, C. (2017) 'Sensitivity to parasitic gaps inside subject islands in native and non-native sentence processing', *Bilingualism: Language and Cognition*, 20(3), pp. 494–511. doi: 10.1017/S1366728915000942.

Braze, D., Shankweiler, D., Ni, W. and Palumbo, L. C. (2002) 'Readers' Eye Movements Distinguish Anomalies of Form and Content', *Journal of Psycholinguistic Research*, 31(1), pp. 25–44. doi: 10.1023/A:1014324220455.

Bresnan, J. (1972) *Theory of complementation in English syntax*. MIT.

Brysbaert, M. (2013) 'Lextale_FR A Fast, Free, and Efficient Test to Measure Language Proficiency in French', *Psychologica Belgica*, 53(1), p. 23. doi: 10.5334/pb-53-1-23.

Brysbaert, M., Drieghe, D. and Vitu, F. (2005) 'Word skipping: implications for theories of eye movement control in reading', in Underwood, G. (ed.) *Cognitive Processes in Eye Guidance*. Oxford: Oxford University Press, pp. 53–78.

Brysbaert, M., Lagrou, E. and Stevens, M. (2017) 'Visual word recognition in a second language:

A test of the lexical entrenchment hypothesis with lexical decision times', *Bilingualism: Language and Cognition*, 20(3), pp. 530–548. doi: 10.1017/S1366728916000353.

Caha, P. (2009) *The nanosyntax of case*. University of Tromsø.

Cameron, R. D. (2011) *Native and nonnative processing of modality and mood in Spanish*. The Florida State University.

Carroll, S. E. (1999) 'Putting "input" in its proper place', *Second Language Research*, 15(4), pp. 337–88.

Carroll, S. E. (2001) *Input and evidence: the raw material of second language acquisition*. Amsterdam: John Benjamins Publishing Company.

Carroll, S. E. (2007) 'Autonomous induction theory', in VanPatten, B. and Williams, J. (eds) *Theories in second language acquisition*. New Jersey: Lawrence Erlbaum Associates, pp. 155–200.

Chambers, C. G. and Cooke, H. (2009) 'Lexical competition during second-language listening: Sentence context, but not proficiency, constrains interference from the native lexicon.', *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35(4), pp. 1029–1040. doi: 10.1037/a0015901.

Chamorro, G., Sorace, A. and Sturt, P. (2016) 'What is the source of L1 attrition? The effect of recent L1 re-exposure on Spanish speakers under L1 attrition', *Bilingualism: Language and Cognition*, 19(3), pp. 520–532. doi: 10.1017/S1366728915000152.

Cho, J. and Slabakova, R. (2014) 'Interpreting definiteness in a second language without articles: The case of L2 Russian', *Second Language Research*, 30(2), pp. 159–190. doi: 10.1177/0267658313509647.

Chomsky, N. (1995) *The minimalist program*. Cambridge, MA: MIT Press.

Chomsky, N. (1998) 'Minimalist Inquiries: The Framework', *MIT Working Papers in Linguistics*, 15, pp. 1–56.

Chomsky, N. (2000) *New horizons in the study of language and mind*. Cambridge: Cambridge University Press.

Christianson, K., Hollingworth, A., Halliwell, J. F. and Ferreira, F. (2001) 'Thematic Roles Assigned along the Garden Path Linger', *Cognitive Psychology*, 42(4), pp. 368–407. doi: 10.1006/cogp.2001.0752.

Clahsen, H., Balkhair, L., Schutter, J.-S. and Cunnings, I. (2013) 'The time course of morphological processing in a second language', *Second Language Research*, 29(1), pp. 7–31. doi: 10.1177/0267658312464970.

Clahsen, H. and Felser, C. (2006) 'Grammatical processing in language learners', *Applied Psycholinguistics*, 27(1), pp. 3–42. doi: 10.1017/S0142716406060024.

Clahsen, H. and Felser, C. (2018) 'Some notes on the shallow structure hypothesis', *Studies in*

Second Language Acquisition, 40(3), pp. 693–706. doi: 10.1017/S0272263117000250.

Clark, R. and Roberts, I. (1993) 'A computational approach to language learnability and language change.', *Linguistic Inquiry*, 24(2), pp. 299–345.

Cleeremans, A. and Jiménez, L. (2002) 'Implicit learning and consciousness: A graded, dynamic perspective', in French, R. M. and Cleeremans, A. (eds) *Implicit learning and consciousness*. Hove: Psychology Press, pp. 1–40.

Colheart, M., Rastle, K., Perry, C., Langdon, R. and Ziegler, J. C. (2001) 'DRC: A dual route cascaded model of visual word recognition and reading aloud', *Psychological Review*, 108, pp. 204–256.

Collentine, J. (2004) 'The Effects of Learning Contexts on Morphosyntactic and Lexical Development', *Studies in Second Language Acquisition*, 26(2), pp. 227–248. doi: 10.1017/S0272263104262040.

Conklin, K. and Pellicer-Sánchez, A. (2016) 'Using eye-tracking in applied linguistics and second language research', *Second Language Research*, 32(3), pp. 453–467. doi: 10.1177/0267658316637401.

Crawford, W. J. (2009) 'The mandative subjunctive', in Rohdenburg, G. and Schlüter, J. (eds) *One Language, Two Grammars?* Cambridge: Cambridge University Press, pp. 257–276. doi: 10.1017/CBO9780511551970.015.

Cunnings, I. (2017) 'Parsing and Working Memory in Bilingual Sentence Processing', *Bilingualism: Language and Cognition*, 20(4), pp. 659–678. doi: 10.1017/S1366728916000675.

Dallas, A. C. (2008) *Influences of verbal properties on second-language filler-gap resolution: A cross-methodological study*. University of Florida, Gainesville, FL.

Dekydtspotter, L. and Renaud, C. (2014) 'On second language processing and grammatical development', *Linguistic Approaches to Bilingualism*, 4(2), pp. 131–165. doi: 10.1075/lab.4.2.01dek.

Dekydtspotter, L., Schwartz, B. D. and Sprouse, R. A. (2006) 'The comparative fallacy in L2 processing research', in O'Brien, M. G., Shea, C., and Archibald, J. (eds) *Proceedings of the 8th Generative Approaches to Second Language Acquisition Conference (GASLA 2006)*. Somerville, MA: Cascadilla Proceedings Project, pp. 33–40.

Dell, G. S. and Chang, F. (2013) 'The P-chain: relating sentence production and its disorders to comprehension and acquisition', *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1634), pp. 20120394–20120394. doi: 10.1098/rstb.2012.0394.

Denison, D. (1998) 'Syntax', in Romaine, S. (ed.) *The Cambridge History of the English Language, vol. 4: 1776-1997*. Cambridge: Cambridge University Press, pp. 92–329.

Dienes, Z. and Perner, J. (1999) 'A theory of implicit and explicit knowledge', *Behavioral and*

Brain Sciences, 22(5), pp. 735–808. doi: 10.1017/S0140525X99002186.

Diependaele, K., Lemhöfer, K. and Brysbaert, M. (2013) 'The word frequency effect in first- and second-language word recognition: A lexical entrenchment account', *Quarterly Journal of Experimental Psychology*, 66(5), pp. 843–863. doi: 10.1080/17470218.2012.720994.

Dijkstra, T. (2005) 'Bilingual word recognition and lexical access', in Kroll, J. F. and De Groot, A. M. . (eds) *Handbook of bilingualism: Psycholinguistic approaches*. New York: Cambridge University Press, pp. 179–201.

Dijkstra, T. and van Heuven, W. J. B. (2002) 'The architecture of the bilingual word recognition system: From identification to decision', *Bilingualism: Language and Cognition*, 5(03). doi: 10.1017/S1366728902003012.

Domínguez, L. (2007) 'Review article: Knowledge of features in fossilized second language grammars', *Second Language Research*, 23(2), pp. 243–260. doi: 10.1177/0267658307076595.

Domínguez, L., Arche, M. J. and Myles, F. (2017) 'Spanish Imperfect revisited: Exploring L1 influence in the reassembly of imperfective features onto new L2 forms', *Second Language Research*, 33(4), pp. 431–457. doi: 10.1177/0267658317701991.

Dowty, D. (1991) 'Thematic proto-roles and argument selection', *Language*, 67(3), pp. 547–619. doi: 10.1353/lan.1991.0021.

Duffy, S. A., Morris, R. K. and Rayner, K. (1988) 'Lexical ambiguity and fixation times in reading', *Journal of Memory and Language*, 27(4), pp. 429–446. doi: 10.1016/0749-596X(88)90066-6.

Dussias, P. E. and Sagarra, N. (2007) 'The effect of exposure on syntactic parsing in Spanish–English bilinguals', *Bilingualism: Language and Cognition*, 10(1), pp. 101–116. doi: 10.1017/S1366728906002847.

Dussias, P. E., Valdés Kroff, J. R., Guzzardo Tamargo, R. E. and Gerfen, C. (2013) 'When gender and looking go hand in hand: Grammatical gender processing in L2 Spanish', *Studies in Second Language Acquisition*, 35(2), pp. 353–387. doi: 10.1017/S0272263112000915.

Dyke, J. A. van, Johns, C. L. and Kukona, A. (2014) 'Low working memory capacity is only spuriously related to poor reading comprehension', *Cognition*, 131(3), pp. 373–403. doi: 10.1016/j.cognition.2014.01.007.

Ehrlich, S. F. and Rayner, K. (1981) 'Contextual effects on word perception and eye movements during reading', *Journal of Verbal Learning and Verbal Behavior*, 20(6), pp. 641–655.

Ellis, N. C. (1994) 'Implicit and explicit processes in language acquisition: An introduction', in Ellis, N. C. (ed.) *Implicit and Explicit Learning of Languages*. San Diego, CA: Academic

Press, pp. 1–32.

Ellis, N. C., Hafeez, K., Martin, K. I., Chen, L., Boland, J. and Sagarra, N. (2014) 'An eye-tracking study of learned attention in second language acquisition', *Applied Psycholinguistics*, 35(3), pp. 547–579. doi: 10.1017/S0142716412000501.

Ellis, R. (2005) 'Measuring implicit and explicit knowledge of a second language: A psychometric study', *Studies in Second Language Acquisition*, 27(2), pp. 141–172. doi: 10.1017/S0272263105050096.

Ellis, R. (2009) 'Implicit and Explicit Learning, Knowledge and Instruction', in Ellis, R., Loewen, S., Elder, C., Reinders, H., Erlam, R., and Philp, J. (eds) *Implicit and Explicit Knowledge in Second Language Learning, Testing and Teaching*. Bristol, Blue Ridge Summit: Multilingual Matters, pp. 3–26. doi: 10.21832/9781847691767-003.

Elman, J. L. L., Bates, E. A. A., Johnson, M., Karmiloff-Smith, A., Parisi, D. and Plunkett, K. (1996) *Rethinking Innateness: A Connectionist Perspective on Development*. Cambridge, MA: MIT Press.

Emonds, J. (1970) *Root and Structure-Preserving Transformations in English*. MIT.

Emonds, J. (1976) *A Transformational Approach to English Syntax*. New York: Academic Press.

Engbert, R., Nuthmann, A., Richter, E. M. and Kliegl, R. (2005) 'SWIFT: A Dynamical Model of Saccade Generation During Reading.', *Psychological Review*, 112(4), pp. 777–813. doi: 10.1037/0033-295X.112.4.777.

Fauconnier, G. (1994) *Mental spaces*. Cambridge, UK: Cambridge University Press.

Felser, C. and Cunnings, I. (2012) 'Processing reflexives in a second language: The timing of structural and discourse-level constraints', *Applied Psycholinguistics*, 33(3), pp. 571–603. doi: 10.1017/S0142716411000488.

Felser, C., Cunnings, I., Batterham, C. and Clahsen, H. (2012) 'The timing of island effects in nonnative sentence processing', *Studies in Second Language Acquisition*, 34(1), pp. 67–98. doi: 10.1017/S0272263111000507.

Felser, C. and Roberts, L. (2007) 'Processing wh-dependencies in a second language: a cross-modal priming study', *Second Language Research*, 23(1), pp. 9–36. doi: 10.1177/0267658307071600.

Felser, C., Sato, M. and Bertenshaw, N. (2009) 'The on-line application of binding Principle A in English as a second language', *Bilingualism: Language and Cognition*, 12(4), pp. 485–502. doi: 10.1017/S1366728909990228.

Ferguson, H. J. (2012) 'Eye movements reveal rapid concurrent access to factual and counterfactual interpretations of the world', *Quarterly Journal of Experimental Psychology*, 65(5), pp. 939–961. doi: 10.1080/17470218.2011.637632.

Ferguson, H. J. and Cane, J. E. (2015) 'Examining the cognitive costs of counterfactual language comprehension: Evidence from ERPs', *Brain Research*, 1622, pp. 252–269. doi: 10.1016/j.brainres.2015.05.024.

Ferguson, H. J. and Sanford, A. J. (2008) 'Anomalies in real and counterfactual worlds: An eye-movement investigation', *Journal of Memory and Language*, 58(3), pp. 609–626. doi: 10.1016/j.jml.2007.06.007.

Ferguson, H. J., Scheepers, C. and Sanford, A. J. (2010) 'Expectations in counterfactual and theory of mind reasoning', *Language and Cognitive Processes*, 25(3), pp. 297–346. doi: 10.1080/01690960903041174.

Ferreira, F., Bailey, K. G. D. and Ferraro, V. (2002) 'Good-Enough Representations in Language Comprehension', *Current Directions in Psychological Science*, 11(1), pp. 11–15. doi: 10.1111/1467-8721.00158.

Ferreira, F. and Chantavarin, S. (2018) 'Integration and Prediction in Language Processing: A Synthesis of Old and New', *Current Directions in Psychological Science*, 27(6), pp. 443–448.

Ferreira, F. and Patson, N. D. (2007) 'The "Good Enough" Approach to Language Comprehension', *Language and Linguistics Compass*, 1(1–2), pp. 71–83. doi: 10.1111/j.1749-818X.2007.00007.x.

Field, A., Miles, J. and Field, Z. (2014) *Discovering Statistics Using R*. Thousand Oaks, CA: SAGE Publications.

Filiaci, F., Sorace, A. and Carreiras, M. (2014) 'Anaphoric biases of null and overt subjects in Italian and Spanish: a cross-linguistic comparison', *Language, Cognition and Neuroscience*, 29(7), pp. 825–843. doi: 10.1080/01690965.2013.801502.

Fintel, K. von (2006) 'Modality and Language', in Borchert., D. M. (ed.) *Encyclopedia of Philosophy - Second Edition*. Detroit: MacMillan Reference USA. Available at: http://web.mit.edu/fintel/fintel-2006-modality.pdf.

Fodor, J. (1983) *The modularity of mind: An essay on faculty psychology*. Cambridge, MA: MIT Press.

Fodor, J. (2000) *The mind doesn't work that way: The scope and limits of computational psychology*. Cambridge, MA: MIT Press.

Fodor, J. D. (1998a) 'Leaning to Parse?', *Journal of Psycholinguistic Research*, 27(2), pp. 285–319. doi: 10.1353/cjl.2004.0023.

Fodor, J. D. (1998b) 'Parsing to Learn', *Journal of Psycholinguistic Research*, 27(3), pp. 339–374. doi: 10.1023/A:1023255705029.

Fodor, J. D. and Sakas, W. (2017) 'Learnability', in Roberts, I. (ed.) *The Oxford Handbook of Universal Grammar*. Oxford: Oxford University Press, pp. 249–269. doi: 10.1093/oxfordhb/9780199573776.013.11.

Foote, R. (2011) 'Integrated knowledge of agreement in early and late English–Spanish bilinguals', *Applied Psycholinguistics*, 32(1), pp. 187–220. doi:

10.1017/S0142716410000342.

Foucart, A. and Frenck-Mestre, C. (2012) 'Can late L2 learners acquire new grammatical features? Evidence from ERPs and eye-tracking', *Journal of Memory and Language*, 66(1), pp. 226–248. doi: 10.1016/j.jml.2011.07.007.

Franceschina, F. (2001) 'Morphological or syntactic deficits in near-native speakers? An assessment of some current proposals', *Second Language Research*, 17(3), pp. 213–247. doi: 10.1177/026765830101700301.

Franceschina, F. (2005) *Fossilized Second Language Grammars*. Amsterdam: John Benjamins Publishing Company (Language Acquisition and Language Disorders). doi: 10.1075/lald.38.

Frazier, L. (1978) *On comprehending sentences: Syntactic parsing strategies*. University of Connecticut, Storrs.

Frazier, L. (1987) 'Sentence processing: a tutorial review', in Coltheart, M. (ed.) *Attention and performance: Vol. 12. The psychology of reading*. Hillsdale, NJ: Erlbaum., pp. 559–586.

Freed, B., Lazar, N. and So, S. (2003) 'Language learning abroad: How do gains -in written fluency compare with oral fluency in French as a second language?', *ADFL Bulletin*, 34, pp. 34–40.

Frenck-Mestre, C. (2002) 'An on-line look at sentence processing in the second language', in Heredia, R. R. and Altarriba, J. (eds) *Advances in psychology: Vol. 134. Bilingual sentence processing*. North-Holland: Elsevier Science Publishers, pp. 217–236. doi: 10.1016/S0166-4115(02)80012-7.

Gelderen, E. van (2004) *Grammaticalizaton as Economy*. Philadephia: John Benjamins Publishing Company.

Gernsbacher, M. A. (1991) 'Cognitive processes and mechanisms in language comprehension: The structure building framework.', in Bower, G. H. (ed.) *The psychology of learning and motivation*. New York, NY: Academic Press, pp. 217–263. doi: 10.1016/S0079-7421(08)60125-5.

Giannakidou, A. (1998) *Polarity Sensitivity as (Non)veridical Dependency*. Amsterdam: John Benjamins Publishing Company.

Giannakidou, A. (2009) 'The dependency of the subjunctive revisited: Temporal semantics and polarity', *Lingua*, 119(12), pp. 1883–1908. doi: 10.1016/j.lingua.2008.11.007.

Gil, K.-H. and Marsden, H. (2013) 'Existential quantifiers in second language acquisition', *Linguistic Approaches to Bilingualism*, 3(2), pp. 117–149. doi: 10.1075/lab.3.2.01gil.

Godfroid, A. (2019) *Eye Tracking in Second Language Acquisition and Bilingualism*. London: Routledge. doi: 10.4324/9781315775616.

Godfroid, A., Loewen, S., Jung, S., Park, J.-H., Gass, S. and Ellis, R. (2015) 'Timed and Untimed

Grammaticality Judgments Measure Distinct Types of Knowledge', *Studies in Second Language Acquisition*, 37(2), pp. 269–297. doi: 10.1017/S0272263114000850.

Godfroid, A. and Winke, P. (2015) 'Investigating implicit and explicit processing using L2 learners' eye-movement data', in Rebuschat, P. (ed.) *Implicit and Explicit Learning of Languages*. Amsterdam: John Benjamins Publishing Company, pp. 323–348. doi: 10.1075/sibil.48.14god.

Gompel, R. P. G. van, Pickering, M. J. and Traxler, M. J. (2000) 'Unrestricted race: A new model of syntactic ambiguity resolution', in Kennedy, A., Radach, R., Heller, D., and Pynte, J. (eds) *Reading as a perceptual process*. Oxford, England: Elsevier, pp. 621–648.

Gompel, R. P. G. van, Pickering, M. J. and Traxler, M. J. (2001) 'Reanalysis in Sentence Processing: Evidence against Current Constraint-Based and Two-Stage Models', *Journal of Memory and Language*, 45(2), pp. 225–258. doi: 10.1006/jmla.2001.2773.

Graesser, A. C. and Forsyth, C. (2013) *Discourse Comprehension*. Oxford University Press. doi: 10.1093/oxfordhb/9780195376746.013.0030.

Green, D. W. and Abutalebi, J. (2013) 'Language control in bilinguals: The adaptive control hypothesis', *Journal of Cognitive Psychology*, 25(5), pp. 515–530. doi: 10.1080/20445911.2013.796377.

Gregg, K. R. (1996) 'The logical and developmental problems of second language acquisition', in Ritchie, W. C. (ed.) *Handbook of second language acquisition*. San Diego, CA: Academic Press, pp. 49–81.

Gregg, K. R. (2003) 'The state of emergentism in second language acquisition', *Second Language Research*, 19(2), pp. 95–128. doi: 10.1191/0267658303sr213oa.

Grey, S., Cox, J. G., Serafini, E. J. and Sanz, C. (2015) 'The Role of Individual Differences in the Study Abroad Context: Cognitive Capacity and Language Development During Short-Term Intensive Language Exposure', *The Modern Language Journal*, 99(1), pp. 137–157. doi: 10.1111/modl.12190.

Grüter, T., Rohde, H. and Schafer, A. J. (2017) 'Coreference and discourse coherence in L2', *Linguistic Approaches to Bilingualism*, 7(2), pp. 199–229. doi: 10.1075/lab.15011.gru.

Guijarro-Fuentes, P. (2012) 'The acquisition of interpretable features in L2 Spanish: Personal a', *Bilingualism: Language and Cognition*, 15(4), pp. 701–720. doi: 10.1017/S1366728912000144.

Gutiérrez, X. (2013) 'The construct validity of grammaticality judgment tests as measures of implicit and explicit knowledge', *Studies in Second Language Acquisition*, 35(3), pp. 423–449. doi: 10.1017/S0272263113000041.

Haigh, M. and Stewart, A. J. (2011) 'The influence of clause order, congruency, and probability on the processing of conditionals', *Thinking & Reasoning*, 17(4), pp. 402–423. doi:

10.1080/10463283.2011.628000.

Hale, J. T. (2011) 'What a Rational Parser Would Do', *Cognitive Science*, 35(3), pp. 399–443. doi: 10.1111/j.1551-6709.2010.01145.x.

Harm, M. W. and Seidenberg, M. S. (2004) 'Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes.', *Psychological Review*, 111(3), pp. 662–720.

Harrington, M. (2001) 'Sentence Processing', in Robinson, P. (ed.) *Cognition and Second Language Instruction*. New York: Cambridge University Press, pp. 70–124.

Harris, M. (1978) *The evolution of French syntax : comparative approach*. London/New York: Longman.

Havik, E., Roberts, L., van Hout, R., Schreuder, R. and Haverkort, M. (2009) 'Processing Subject-Object Ambiguities in the L2: A Self-Paced Reading Study With German L2 Learners of Dutch', *Language Learning*, 59(1), pp. 73–112. doi: 10.1111/j.1467-9922.2009.00501.x.

Hawkins, R. and Chan, C. Y.-H. (1997) 'The partial availability of Universal Grammar in second language acquisition: the "failed functional features hypothesis"', *Second Language Research*, 13(3), pp. 187–226. doi: 10.1191/026765897671476153.

Hawkins, R. and Franceschina, F. (2004) 'Explaining the acquisition and non-acquisition of determiner-noun gender concord in French and Spanish', in Prévost, P. and Paradis, J. (eds) *The Acquisition of French in Different Contexts: Focus on functional categories*. Amsterdam: John Benjamins Publishing Company, pp. 175–205. doi: 10.1075/lald.32.10haw.

Hawkins, R. and Hattori, H. (2006) 'Interpretation of English multiple wh-questions by Japanese speakers: a missing uninterpretable feature account', *Second Language Research*, 22(3), pp. 269–301. doi: 10.1191/0267658306sr269oa.

Hawkins, R. and Towell, R. (2015) *French Grammar and Usage*. Milton Park, Oxfordshire: Routledge.

Haznedar, B. and Schwartz, B. D. (1997) 'Are there optional infinitives in child L2 acquisition?', in Hughes, E., Hughes, M., and Greenhill, A. (eds) *Proceedings of the 21st Annual Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press., pp. 257–268.

Henderson, J. M. and Luke, S. G. (2012) 'Oculomotor inhibition of return in normal and mindless reading', *Psychonomic Bulletin & Review*, 19(6), pp. 1101–1107. doi: 10.3758/s13423-012-0274-2.

Hopp, H. (2006) 'Syntactic features and reanalysis in near-native processing', *Second Language Research*, 22(3), pp. 369–397. doi: 10.1191/0267658306sr272oa.

Hopp, H. (2007) *Ultimate attainment at the interfaces in second language acquisition: Grammar and* *processing*. University of Groningen, Groningen, Netherlands.

Hopp, H. (2010) 'Ultimate attainment in L2 inflection: Performance similarities between non-native and native speakers', *Lingua*, 120(4), pp. 901–931. doi: 10.1016/j.lingua.2009.06.004.

Hopp, H. (2013) 'Grammatical gender in adult L2 acquisition: Relations between lexical and syntactic variability', *Second Language Research*, 29(1), pp. 33–56. doi: 10.1177/0267658312461803.

Hopp, H. (2015) 'Individual differences in the second language processing of object–subject ambiguities', *Applied Psycholinguistics*, 36(2), pp. 129–173. doi: 10.1017/S0142716413000180.

Hopp, H. (2017a) 'Cross-linguistic lexical and syntactic co-activation in L2 sentence processing', *Linguistic Approaches to Bilingualism*, 7(1), pp. 96–130. doi: 10.1075/lab.14027.hop.

Hopp, H. (2017b) 'Individual differences in L2 parsing and lexical representations', *Bilingualism: Language and Cognition*, 20(4), pp. 689–690. doi: 10.1017/S1366728916000821.

Hopp, H., Bail, J. and Jackson, C. N. (2020) 'Frequency at the syntax–discourse interface: A bidirectional study on fronting options in L1/L2 German and L1/L2 English', *Second Language Research*, 36(1), pp. 65–96. doi: 10.1177/0267658318802985.

Hopp, H. and León Arriaga, M. E. (2016) 'Structural and inherent case in the non-native processing of Spanish: Constraints on inflectional variability', *Second Language Research*, 32(1), pp. 75–108. doi: 10.1177/0267658315605872.

Howard, M. (2005a) 'Second language acquisition in a study abroad context: A comparative investigation of the effects of study abroad and foreign language instruction on the L2 learner's grammatical development.', in Housen, A. and Pierrard, M. (eds) *Investigations in instructed second language acquisition*. Berlin, German: Mouton de Gruyter, pp. 495–530.

Howard, M. (2005b) 'The emergence and use of the "plus-que-parfait" in advanced French interlanguage', in Dewaele, J.-M. (ed.) *Focus on French as a foreign language: Multidisciplinary perspectives*. Clevedon, England: Multilingual Matters, pp. 63–87.

Howard, M. (2008) 'Morpho-syntactic development in the expression of modality : The subjunctive in French L2 acquisition', *Canadian Journal of Applied Linguistics*, 11(3), pp. 171–192.

Howard, M. (2012a) 'From tense and aspect towards modality—The acquisition of future, conditional and subjunctive morphology in L2 French. A preliminary study', in Labeau, E. and Saddour, I. (eds) *Tense, aspect, and mood in first and second language acquisition*. Amsterdam, The Netherlands/New York, NY: Rodopi, pp. 201–223.

Howard, M. (2012b) 'The Advanced Learner's Sociolinguistic Profile: On Issues of Individual Differences, Second Language Exposure Conditions, and Type of Sociolinguistic

Variable', *The Modern Language Journal*, 96(1), pp. 20–33. doi: 10.1111/j.1540-4781.2012.01293.x.

Howard, M. and Schwieter, J. W. (2018) 'The Development of Second Language Grammar in a Study Abroad Context', in Cristina Sanz, A. M.-F. (ed.) *The Routledge Handbook of Study Abroad Research and Practice*. Milton Park, Oxfordshire: Routledge, pp. 134–148. doi: 10.4324/9781315639970-9.

Hundt, M. (1998) 'It is important that this study (should) be based on the analysis of parallel corpora: on the use of the mandative subjunctive in four varieties of English', in Lindquist, H., Klintborg, S., Levin, M., and Estling, M. (eds) *The Major Varieties of English*. Växjö: Växjö University, pp. 159–75.

Hundt, M. (2018) 'It is time that this (should) be studied across a broader range of Englishes', in Deshors, S. C. (ed.) *Modeling World Englishes: Assessing the interplay of emancipation and globalization of ESL varieties*. Amsterdam: John Benjamins Publishing Company, pp. 217–244. doi: 10.1075/veaw.g61.09hun.

Hundt, M. and Gardner, A. C. (2017) 'Corpus-based approaches: Watching English change', in Brinton, L. (ed.) *English Historical Linguistics: Approaches and Perspectives*. Cambridge: Cambridge University Press, pp. 96–130.

Hwang, S. H. and Lardiere, D. (2013) 'Plural-marking in L2 Korean: A feature-based approach', *Second Language Research*, 29(1), pp. 57–86. doi: 10.1177/0267658312461496.

Inhoff, A. W. (1984) 'Two stages of word processing during eye fixations in the reading of prose', *Journal of Verbal Learning and Verbal Behavior*, 23(5), pp. 612–624. doi: 10.1016/S0022-5371(84)90382-7.

Inhoff, A. W. and Rayner, K. (1986) 'Parafoveal word processing during eye fixations in reading: Effects of word frequency', *Perception & Psychophysics*, 40(6), pp. 431–440.

Ionin, T. (2011) 'Formal Theory-Based Methodologies', in Mackey, A. and Gass, S. M. (eds) *Research Methods in Second Language Acquisition*. Hoboken, New Jersey: Blackwell Publishing Ltd, pp. 30–52.

Irwin, D. E. (1998) 'Lexical Processing during Saccadic Eye Movements', *Cognitive Psychology*, 36(1), pp. 1–27. doi: 10.1006/cogp.1998.0682.

Irwin, D. E. and Carlson-Radvansky, L. A. (1996) 'Cognitive Suppression During Saccadic Eye Movements', *Psychological Science*, 7(2), pp. 83–88. doi: 10.1111/j.1467-9280.1996.tb00334.x.

Isabelli-García, C. (2010) 'Acquisition of Spanish Gender Agreement in Two Learning Contexts: Study Abroad and At Home', *Foreign Language Annals*, 43(2), pp. 289–303. doi: 10.1111/j.1944-9720.2010.01079.x.

Ivanov, I. P. (2012) 'L2 acquisition of Bulgarian clitic doubling: A test case for the Interface

Hypothesis', *Second Language Research*, 28(3), pp. 345–368. doi: 10.1177/0267658312452066.

Iverson, M., Kempchinsky, P. and Rothman, J. (2008) 'Interface vulnerability and knowledge of the subjunctive / indicative distinction with negated epistemic predicates in L2 Spanish', *EUROSLA Yearbook*, 8, pp. 135–163. doi: 10.1075/eurosla.8.09ive.

Jackson, C. (2008) 'Proficiency Level and the Interaction of Lexical and Morphosyntactic Information During L2 Sentence Processing', *Language Learning*, 58(4), pp. 875–909. doi: 10.1111/j.1467-9922.2008.00481.x.

Jackson, C. N. and Bobb, S. C. (2009) 'The processing and comprehension of wh-questions among second language speakers of German', *Applied Psycholinguistics*, 30(4), pp. 603–636. doi: 10.1017/S014271640999004X.

Jacob, G. and Felser, C. (2016) 'Reanalysis and semantic persistence in native and non-native garden-path recovery', *Quarterly Journal of Experimental Psychology*, 69(5), pp. 907–925. doi: 10.1080/17470218.2014.984231.

Jegerski, J. and Sekerina, I. A. (2020) 'The processing of input with differential object marking by heritage Spanish speakers', *Bilingualism: Language and Cognition*, 23(2), pp. 274–282. doi: 10.1017/S1366728919000087.

Jeppesen Kragh, K. (2010) *Le remplacement de l'imparfait du subjonctif par le présent du subjonctif considéré dans une perspective de grammaticalisation*. Copenhagen: Museum Tusculanum Press.

Jiang, N. (2004) 'Morphological insensitivity in second language processing', *Applied Psycholinguistics*, 25(4), pp. 603–634. doi: 10.1017/S0142716404001298.

Jiang, N. (2007) 'Selective Integration of Linguistic Knowledge in Adult Second Language Learning', *Language Learning*, 57(1), pp. 1–33. doi: 10.1111/j.1467-9922.2007.00397.x.

Jiang, N., Novokshanova, E., Masuda, K. and Wang, X. (2011) 'Morphological Congruency and the Acquisition of L2 Morphemes', *Language Learning*, 61(3), pp. 940–967. doi: 10.1111/j.1467-9922.2010.00627.x.

Johansson, S. and Norheim, E. H. (1988) 'The subjunctive in British and American English', *ICAME Journal*, 12, pp. 27–36.

Johnson-Laird, P. N. and Byrne, R. M. J. (2002) 'Conditionals: A theory of meaning, pragmatics, and inference.', *Psychological Review*, 109(4), pp. 646–678. doi: 10.1037/0033-295X.109.4.646.

Juffs, A. (2001) 'Psycholinguistically oriented second language research', *Annual Review of Applied Linguistics*, 21, pp. 207–220. doi: 10.1017/S0267190501000125.

Juhasz, B. J. (2008) 'The processing of compound words in English: Effects of word length on eye movements during reading', *Language and Cognitive Processes*, 23(7–8), pp. 1057–1088.

Juhasz, B. J., Inhoff, A. W. and Rayner, K. (2005) 'The role of interword spaces in the processing of English compound words', *Language and Cognitive Processes*, 20(1–2), pp. 291–316.

Juhasz, B. J., White, S. J., Liversedge, S. P. and Rayner, K. (2008) 'Eye movements and the use of parafoveal word length information in reading', *Journal of Experimental Psychology*, 34(6), pp. 1560–1579.

Just, M. A. and Carpenter, P. A. (1980) 'A theory of reading: From eye fixations to comprehension.', *Psychological Review*, 87(4), pp. 329–354. doi: 10.1037/0033-295X.87.4.329.

Kaan, E. (2014) 'Predictive sentence processing in L2 and L1', *Linguistic Approaches to Bilingualism*, 4(2), pp. 257–282. doi: 10.1075/lab.4.2.05kaa.

Kaan, E., Dallas, A. C. and Wijnen, F. (2010) 'Syntactic predictions in second-language sentence processing.', in Zwart, J.-W. and de Vries, M. (eds) *Structure preserved: Festschrift in the honor of Jan Koster*. Amsterdam: John Benjamins Publishing Company, pp. 207–2013.

Kastronic, L. (2016) *A Comparative Variationist Approach to Morphosyntactic Variation in Hexagonal and Quebec French*. University of Ottawa.

Keating, G. D. (2009) 'Sensitivity to Violations of Gender Agreement in Native and Nonnative Spanish: An Eye-Movement Investigation', *Language Learning*, 59(3), pp. 503–535. doi: 10.1111/j.1467-9922.2009.00516.x.

Keating, G. D. and Jegerski, J. (2015) 'Experimental Designs in Sentence Processing Research', *Studies in Second Language Acquisition*, 37(1), pp. 1–32. doi: 10.1017/S0272263114000187.

Kempchinsky, P. (2009) 'What can the subjunctive disjoint reference effect tell us about the subjunctive?', *Lingua*, 119(12), pp. 1788–1810. doi: 10.1016/j.lingua.2008.11.009.

Kintsch, W. (1991) 'The Role of Knowledge in Discourse Comprehension: A Construction-Integration Model', *Psychological Review*, 95(2), pp. 163–182. doi: 10.1037/0033-295X.95.2.163.

Kintsch, W. (1998) *Comprehension: A paradigm for cognition*. New York, NY, US: Cambridge University Press.

Kintsch, W. and van Dijk, T. A. (1978) 'Toward a model of text comprehension and production.', *Psychological Review*, 85(5), pp. 363–394. doi: 10.1037/0033-295X.85.5.363.

Krashen, S. (1981) *Second language acquisition and second language learning*. Oxford: Pergamon.

Kroll, J. F. and Bialystok, E. (2013) 'Understanding the consequences of bilingualism for language processing and cognition', *Journal of Cognitive Psychology*, 25(5), pp. 497–514. doi: 10.1080/20445911.2013.799170.

Kroll, J. F., Bobb, S. C. and Wodniecka, Z. (2006) 'Language selectivity is the exception, not the rule: Arguments against a fixed locus of language selection in bilingual speech', *Bilingualism: Language and Cognition*, 9(2), pp. 119–135. doi: 10.1017/S1366728906002483.

Kulakova, E. and Nieuwland, M. S. (2016a) 'Pragmatic skills predict online counterfactual comprehension: Evidence from the N400', *Cognitive, Affective and Behavioral Neuroscience*, 16(5), pp. 814–824. doi: 10.3758/s13415-016-0433-4.

Kulakova, E. and Nieuwland, M. S. (2016b) 'Understanding Counterfactuality: A Review of Experimental Evidence for the Dual Meaning of Counterfactuals', *Language and Linguistics Compass*, 10(2), pp. 49–65. doi: 10.1111/lnc3.12175.

LaBrozzi, R. (2012) 'The role of study abroad and inhibitory control on processing redundant cues', in Geeslin, K. and Díaz-Campos, M. (eds) *Selected proceedings of the 14th Hispanic linguistic symposium*. Somerville, MA: Cascadilla Proceedings Project, pp. 228–241.

Lardiere, D. (2006) 'Attainment and acquirability in second language acquisition', *Second Language Research*, 22(3), pp. 239–242. doi: 10.1191/0267658306sr267ed.

Lardiere, D. (2009) 'Some thoughts on the contrastive analysis of features in second language acquisition', *Second Language Research*, 25(2), pp. 173–227. doi: 10.1177/0267658308100283.

Lasnik, H. (1995) 'Verbal morphology: Syntactic structures meets the Minimalist Program', in Campos, H. and Kempchinksy, P. (eds) *Evolution and Revolution in Linguistic Theory*. Washington, DC: Georgetown University Press, pp. 251–275.

Lasnik, H. and Lidz, J. L. (2016) 'The Argument from the Poverty of the Stimulus', in Roberts, I. (ed.) *The Oxford Handbook of Universal Grammar*. Oxford University Press, pp. 220–248. doi: 10.1093/oxfordhb/9780199573776.013.10.

Leal Méndez, T. and Slabakova, R. (2014) 'The Interpretability Hypothesis again: A partial replication of Tsimpli and Dimitrakopoulou (2007)', *International Journal of Bilingualism*, 18(6), pp. 537–557. doi: 10.1177/1367006912448125.

Leech, G., Hundt, M., Mair, C. and Smith, N. (2009) *Change in Contemporary English*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511642210.

Leeser, M. J., Brandl, A. and Weissglass, C. (2011) 'Task effects in second language sentence processing research', in Trofimovich, P. and McDonough, K. (eds) *Applying priming methods to L2 learning, teaching and research: Insights from Psycholinguistics*. Amsterdam: John Benjamins Publishing Company, pp. 179–198. doi: 10.1075/lllt.30.13lee.

Lemhöfer, K. and Broersma, M. (2012) 'Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English', *Behavior Research Methods*, 44(2), pp. 325–343. doi: 10.3758/s13428-011-0146-0.

Lenoble-Pinson, M. and Grevisse, M. (2009) *Le français correct : Guide pratique des difficultés*. Paris: Duculot.

Lenth, R. (2018) 'emmeans: Estimated marginal means, aka leastsquares means (v. 1.1.2).'

Levitt, J. (1967) 'The subjunctive in Modern French: Conflicts between theory and usage', *Linguistics*, 5(31). doi: 10.1515/ling.1967.5.31.50.

Levy, R. (2008) 'Expectation-based syntactic comprehension', *Cognition*, 106(3), pp. 1126–1177. doi: 10.1016/j.cognition.2007.05.006.

Lewis, R. L. and Vasishth, S. (2005) 'An Activation-Based Model of Sentence Processing as Skilled Memory Retrieval', *Cognitive Science*, 29(3), pp. 375–419. doi: 10.1207/s15516709cog0000_25.

Li, P., Zhang, F., Yu, A. and Zhao, X. (2019) 'Language History Questionnaire (LHQ3): An enhanced tool for assessing multilingual experience', *Bilingualism: Language and Cognition*, pp. 1–7. doi: 10.1017/S1366728918001153.

Libben, M. R. and Titone, D. A. (2009) 'Bilingual lexical access in context: Evidence from eye movements during reading.', *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35(2), pp. 381–390. doi: 10.1037/a0014875.

Lightfoot, D. (1999) *The Development of Language: Acquisition, Change and Evolution*. Cambridge, MA: Wiley-Blackwell.

Lim, J. H. and Christianson, K. (2013a) 'Integrating meaning and structure in L1-L2 and L2-L1 translations', *Second Language Research*, 29(3), pp. 233–256. doi: 10.1177/0267658312462019.

Lim, J. H. and Christianson, K. (2013b) 'Second language sentence processing in reading for comprehension and translation', *Bilingualism*, 16(3), pp. 518–537. doi: 10.1017/S1366728912000351.

Lim, J. H. and Christianson, K. (2015) 'Second language sensitivity to agreement errors: Evidence from eye movements during comprehension and translation', *Applied Psycholinguistics*, 36(6), pp. 1283–1315. doi: 10.1017/S0142716414000290.

Liversedge, S. P. and Findlay, J. M. (2000) 'Saccadic eye movements and cognition', *Trends in Cognitive Sciences*, 4(1), pp. 6–14. doi: 10.1016/S1364-6613(99)01418-7.

Liversedge, S. P., Paterson, K. B. and Pickering, M. J. (1998) 'Eye Movements and Measures of Reading Time', in Underwood, G. (ed.) *Eye Guidance in Reading and Scene Perception*. Amsterdam: Elsevier, pp. 55–75. doi: 10.1016/B978-008043361-5/50004-3.

Lodge, A. (1993) *French: From dialect to standard*. London: Routledge.

Love, T., Maas, E. and Swinney, D. (2003) 'The Influence of Language Exposure on Lexical and Syntactic Language Processing', *Experimental Psychology*, 50(3), pp. 204–216. doi: 10.1026//1617-3169.50.3.204.

MacDonald, M. C. M. C., Pearlmutter, N. J. N. J. and Seidenberg, M. S. M. S. (1994) 'The lexical nature of syntactic ambiguity resolution', *Psychological Review*, 101(4), pp. 673–703. doi: 10.1037/0033-295X.101.4.676.

MacWhinney, B. (2011) 'A unified model', in Robinson, P. and Ellis, N. C. (eds) *Handbook of cognitive linguistics and second language acquisition*. New York, NY, US: Routledge/Taylor &

Francis Group., pp. 341–371.

Mailhac, J.-P. (2000) 'Sens, choix et subjonctif', *Journal of French Language Studies*, 10(2), pp. 229–244. doi: 10.1017/S0959269500000235.

Marian, V. and Spivey, M. (2003) 'Competing activation in bilingual language processing: Within- and between-language competition', *Bilingualism: Language and Cognition*, 6(2), pp. 97–115. doi: 10.1017/S1366728903001068.

Marinis, T., Roberts, L., Felser, C. and Clahsen, H. (2005a) 'Gaps in Second Language Processing', *Studies in Second Language Acquisition*, 27(1), pp. 53–78. doi: 10.1017/S0272263105050035.

Marinis, T., Roberts, L., Felser, C. and Clahsen, H. (2005b) 'Gaps in second language sentence processing', *Studies in Second Language Acquisition*, 27(1), pp. 53–78. doi: 10.1017/S0272263105050035.

Matin, E. (1974) 'Saccadic suppression: A review and an analysis.', *Psychological Bulletin*, 81(12), pp. 899–917. doi: 10.1037/h0037368.

McClelland, J. L. and Rumelhart, D. E. (1981) 'An interactive activation model of context effects in letter perception: Part 1. An account of basic findings.', *Psychological Review*, 88, pp. 375–407.

McDonald, J. L. (2000) 'Grammaticality judgments in a second language: Influences of age of acquisition and native language', *Applied Psycholinguistics*, 21(3), pp. 395–423. doi: 10.1017/S0142716400003064.

McDonald, J. L. (2006) 'Beyond the critical period: Processing-based explanations for poor grammaticality judgment performance by late second language learners', *Journal of Memory and Language*, 55(3), pp. 381–401. doi: 10.1016/j.jml.2006.06.006.

McManus, K. and Mitchell, R. (2015) 'Subjunctive use and development in L2 French: A longitudinal study', *Language, Interaction and Acquisition*, 6(1), pp. 42–73. doi: 10.1075/lia.6.1.02mcm.

Mcmanus, K., Tracy-Ventura, N., Mitchell, R., Richard, L. and Romero de Mills, P. (2014) 'Exploring the Acquisition of the French Subjunctive : Local Syntactic Context or Oral Proficiency ? Mood in French', in Leclercq, P., Edmonds, A., and Hilton, H. (eds) *Measuring L2 Proficiency: Perspectives from SLA*. Bristol: Multilingual Matters, pp. 167–190.

McRae, K., Spivey-Knowlton, M. J. and Tanenhaus, M. K. (1998) 'Modeling the Influence of Thematic Fit (and Other Constraints) in On-line Sentence Comprehension', *Journal of Memory and Language*, 38(3), pp. 283–312. doi: 10.1006/jmla.1997.2543.

Mitchell, R., Tracy-Ventura, N. and Mcmanus, K. (2017) *Anglophone Students Abroad: Identity, social relationships and language learning*. New York, NY, US: Routledge.

Mitchell, R., Tracy-Ventura, N. and McManus, K. (2017) *Anglophone Students Abroad*. New York:

Routledge. doi: 10.4324/9781315194851.

Morrison, R. E. (1984) 'Manipulation of stimulus onset delay in reading: Evidence for parallel programming of saccades.', *Journal of Experimental Psychology: Human Perception and Performance*, 10(5), pp. 667–682. doi: 10.1037/0096-1523.10.5.667.

Mueller, J. L., Hahne, A., Fujii, Y. and Friederici, A. D. (2005) 'Native and Nonnative Speakers' Processing of a Miniature Version of Japanese as Revealed by ERPs', *Journal of Cognitive Neuroscience*, 17(8), pp. 1229–1244. doi: 10.1162/0898929055002463.

Mulder, K. and Hulstijn, J. H. (2011) 'Linguistic Skills of Adult Native Speakers, as a Function of Age and Level of Education', *Applied Linguistics*, 32(5), pp. 475–494. doi: 10.1093/applin/amr016.

Müller, B. and Elsass, A. (1985) 'Le français d'aujourd'hui'. Paris: Klincksieck. Available at: http://catalog.hathitrust.org/api/volumes/oclc/13414526.html.

Ni, W., Fodor, J. D., Crain, S. and Shankweiler, D. (1998) 'Anomaly Detection: Eye Movement Patterns', *Journal of Psycholinguistic Research*, 27(5), pp. 515–539. doi: 10.1023/A:1024996828734.

Nieuwland, M. S. (2013) '"If a lion could speak …": Online sensitivity to propositional truth-value of unrealistic counterfactual sentences', *Journal of Memory and Language*, 68(1), pp. 54–67. doi: 10.1016/j.jml.2012.08.003.

Nieuwland, M. S. and Martin, A. E. (2012) 'If the real world were irrelevant, so to speak: The role of propositional truth-value in counterfactual sentence comprehension', *Cognition*, 122(1), pp. 102–109. doi: 10.1016/j.cognition.2011.09.001.

Nuthmann, A., Engbert, R. and Kliegl, R. (2007) 'The IOVP effect in mindless reading: Experiment and modeling', *Vision Research*, 47(7), pp. 990–1002. doi: 10.1016/j.visres.2006.11.005.

O'Connor DiVito, N. (1997) *Patterns across spoken and written French: Empirical research on the interaction among forms, functions and genres*. Boston: Houghton Mifflin.

O'Grady, W. (2005) *Syntactic carpentry: An emergentist approach to syntax*. Mahwah, NJ: Lawrence Erhlbaum Associates.

O'Regan, J. K. (1990) 'Eye movements and reading', in Kowler, E. (ed.) *Eye movements and their role in visual and cognitive processing: Reviews of oculomotor research*. Amsterdam: Elsevier, pp. 395–453.

O'Regan, J. K. (1992) 'No TitleOptimal viewing position in words and the strategy–tactics theory of eye movements in reading', in Rayner, K. (ed.) *Eye movements and visual cognition: Scene perception and reading*. New York: Springer, pp. 333–354.

O'Regan, J. K. and Levy-Schoen, A. (1987) 'Eye movement strategy and tactics in word recognition and reading', in Coltheart, M. (ed.) *Attention and performance XII: The psychology* *of reading*. Hillsdale, NJ: Erlbaum., pp. 363–383.

Osch, B. van, Hulk, A., Aalberse, S. and Sleeman, P. (2018) 'Implicit and Explicit Knowledge of a Multiple Interface Phenomenon: Differential Task Effects in Heritage Speakers and L2 Speakers of Spanish in The Netherlands', *Languages*, 3(3), p. 25. doi: 10.3390/languages3030025.

Övergaard, G. (1995) *The Mandative Subjunctive in American and British English in the 20th Century*. Uppsala: Ubsaliensis S. Academiae.

Pakulak, E. and Neville, H. J. (2010) 'Proficiency Differences in Syntactic Processing of Monolingual Native Speakers Indexed by Event-related Potentials', *Journal of Cognitive Neuroscience*, 22(12), pp. 2728–2744. doi: 10.1162/jocn.2009.21393.

Pan, H.-Y. and Felser, C. (2011) 'Referential context effects in L2 ambiguity resolution: Evidence from self-paced reading', *Lingua*, 121(2), pp. 221–236. doi: 10.1016/j.lingua.2010.08.003.

Pan, H.-Y., Schimke, S. and Felser, C. (2015) 'Referential context effects in non-native relative clause ambiguity resolution', *International Journal of Bilingualism*, 19(3), pp. 298–313. doi: 10.1177/1367006913515769.

Papadopoulou, D. and Clahsen, H. (2003) 'Parsing Strategies in L1 and L2 Sentence Processing', *Studies in Second Language Acquisition*, 25(4), pp. 501–528. doi: 10.1017/S0272263103000214.

Paradis, M. (2009) *Declarative and Procedural Determinants of Second Languages*. Amsterdam: John Benjamins Publishing Company (Studies in Bilingualism). doi: 10.1075/sibil.40.

Pearlmutter, N. J., Garnsey, S. M. and Bock, K. (1999) 'Agreement Processes in Sentence Comprehension', *Journal of Memory and Language*, 41(3), pp. 427–456. doi: 10.1006/jmla.1999.2653.

Perfetti, C. (2007) 'Reading Ability: Lexical Quality to Comprehension', *Scientific Studies of Reading*, 11(4), pp. 357–383. doi: 10.1080/10888430701530730.

Perfetti, C. A. and Hart, L. (2001) 'The lexical basis of comprehension skill.', in *On the consequences of meaning selection: Perspectives on resolving lexical ambiguity*. Washington: American Psychological Association, pp. 67–86. doi: 10.1037/10459-004.

Perry, C., Ziegler, J. C. and Zorzi, M. (2007) 'Nested incremental modeling in the development of computational theories: The CDP+ model of reading aloud', *Psychological Review*, 114, pp. 273–315.

Phillips, C. and Ehrenhofer, L. (2015) 'The role of language processing in language acquisition', *Linguistic Approaches to Bilingualism*, 5(4), pp. 409–453. doi: 10.1075/lab.5.4.01phi.

Pickering, M. J. and Gambi, C. (2018) 'Predicting while comprehending language: A theory and review', *Psychological Bulletin*, 144(10), pp. 1002–1044. doi: 10.1037/bul0000158.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S. and Patterson, K. (1996) 'Understanding normal and impaired word reading: Computational principles in quasi-regular domains.', *Psychological Review*, 103(1), pp. 56-115. doi: http://dx.doi.org/10.1037/0033-295X.103.1.56.

Poplack, S. (1989) 'The care and handling of a megacorpus: The Ottawa-Hull French project.', in Fasold, R. and Schiffrin, D. (eds) *Language change and variation*. Amsterdam: John Benjamins Publishing Company, pp. 411–451.

Poplack, S. (1990) 'Prescription, intuition et usage: le subjonctif français et la variabilité inhérente', *Langage et société*, 54, pp. 5–33.

Poplack, S., Lealess, A. and Dion, N. (2013) 'The evolving grammar of the French subjunctive', *Probus*, 25(1), pp. 139–195. doi: https://doi.org/10.1515/probus-2013-0005.

Portner, P. (2018) *Mood*. Oxford: Oxford University Press.

Potsdam, E. (1996) *Syntactic Issues in the English Imperative*. University of California, Santa Cruz.

Pozzan, L. and Trueswell, J. C. (2016) 'Second language processing and revision of garden-path sentences: a visual word study', *Bilingualism: Language and Cognition*, 19(3), pp. 636–643. doi: 10.1017/S1366728915000838.

Prévost, P. and White, L. (2000) 'Accounting for morphological variation in second language acquisition: Truncation of missing inflection?', in Friedemann, M. –A. and Rizzi, L. (eds) *The acquisition of syntax*. London: Longman, pp. 202–235.

Quer, J. (1998) *No Mood at the Interface*. The Hague: Holland Academic Graphics.

Quer, J. (2001) 'Interpreting mood', *Probus*, 13(1), pp. 81–111. doi: 10.1515/prbs.13.1.81.

Quer, J. (2009) 'Twists of mood: The distribution and interpretation of indicative and subjunctive', *Lingua*, 119(12), pp. 1779–1787. doi: 10.1016/j.lingua.2008.12.003.

Radford, A. (2007) 'Feature correlations in nominative Case-marking', in Liceras, J., Zobl, H., and Goodluck, H. (eds) *The Role of Formal Features in Second Language Acquisition*. New York: Lawrence Erhlbaum Associates, pp. 83–104.

Ramchand, G. C. (2008a) 'Lexical Items in Complex Predications: Selection as Underassociation', *Nordlyd, Tromsø Working Papers in Linguistics*. Tromsø, 35, pp. 115–141.

Ramchand, G. C. (2008b) *Verb Meaning and the Lexicon*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511486319.

Rankin, T. and Unsworth, S. (2016) 'Beyond poverty: Engaging with input in generative SLA', *Second Language Research*, 32(4), pp. 563–572. doi: 10.1177/0267658316648732.

Rayner, K. (1998) 'Eye movements in reading and information processing: 20 years of research.', *Psychological Bulletin*, 124(3), pp. 372–422. doi: 10.1037/0033-2909.124.3.372.

Rayner, K. (2009) 'Eye movements and attention in reading, scene perception, and visual search', *Quarterly Journal of Experimental Psychology*, 62(8), pp. 1457–1506. doi: 10.1080/17470210902816461.

Rayner, K. and Duffy, S. A. (1986) 'Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity', *Memory & Cognition*, 14, pp. 191–201.

Rayner, K. and Pollatsek, A. (1989) *The psychology of reading*. Englewood Cliffs, NJ, US: Prentice-Hall, Inc.

Rayner, K. and Pollatsek, A. (2006) 'Eye-movement control in reading', in Traxler, M. and Gernsbacher, M. (eds) *Handbook of Psycholinguistics*. Amsterdam: Elsevier, pp. 613–58.

Rayner, K., Warren, T., Juhasz, B. J. and Liversedge, S. P. (2004) 'The Effect of Plausibility on Eye Movements in Reading.', *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30(6), pp. 1290–1301. doi: 10.1037/0278-7393.30.6.1290.

Rayner, K. and Well, A. D. (1996) 'Effects of contextual constraint on eye movements in reading: A further examination', *Psychonomic Bulletin and Review*, 3, pp. 504–509.

Reber, A. S. (1976) 'Implicit learning of synthetic languages: The role of instructional set.', *Journal of Experimental Psychology: Human Learning and Memory*, 2(1), pp. 88–94. doi: 10.1037/0278-7393.2.1.88.

Rebuschat, P. (2013) 'Measuring Implicit and Explicit Knowledge in Second Language Research', *Language Learning*, 63(3), pp. 595–626. doi: 10.1111/lang.12010.

Reichle, E. D., Pollatsek, A., Fisher, D. L. and Rayner, K. (1998) 'Toward a model of eye movement control in reading.', *Psychological Review*, 105(1), pp. 125–157. doi: 10.1037/0033-295X.105.1.125.

Reichle, E. D., Pollatsek, A. and Rayner, K. (2012) 'Using E-Z Reader to simulate eye movements in nonreading tasks: A unified framework for understanding the eye–mind link.', *Psychological Review*, 119(1), pp. 155–185. doi: 10.1037/a0026473.

Rizzi, L. and Bocci, G. (2017) 'Left Periphery of the Clause: Primarily Illustrated for Italian', in Everaert, M. and Riemsdijk, H. van (eds) *The Wiley Blackwell Companion to Syntax, Second Edition*. Hoboken, NJ, USA: John Wiley & Sons, Inc., pp. 1–30. doi: 10.1002/9781118358733.wbsyncom104.

Roberts, I. (1985) 'Agreement parameters and the development of English modal auxiliaries', *Natural Language and Linguistic Theory*, 3, pp. 21–58.

Roberts, I. (1993) *Verbs and Diachrnoic Syntax*. Dordrecht: Kluwer.

Roberts, L. (2018) 'Psycholinguistic Approaches and Advanced Proficiency', in Malovrh, P. A. and Benati, A. G. (eds) *The Handbook of Advanced Proficiency in Second Language Acquisition*. Hoboken, NJ, USA: John Wiley & Sons, Inc., pp. 30–50. doi: 10.1002/9781119261650.ch3.

Roberts, L. (2019) 'Psycholinguistic and Neurolinguistic Methods', in Schweiter, J. and Benati, A. G. (eds) *Cambridge Handbook of Language Learning*. Cambridge: Cambridge Univrsity Press.

Roberts, L. and Felser, C. (2011) 'Plausibility and recovery from garden paths in second language sentence processing', *Applied Psycholinguistics*, 32(2), pp. 299–331. doi: 10.1017/S0142716410000421.

Roberts, L., Gullberg, M. and Indefrey, P. (2008) 'Online pronoun resolution in L2 discourse: L1 Influence and General Learner Effects', *Studies in Second Language Acquisition*, 30(3), pp. 333–357. doi: 10.1017/S0272263108080480.

Roberts, L., Sagarra, N. and Herschensohn, J. (2013) 'Processing of gender and number agreement in late Spanish bilinguals', *International Journal of Bilingualism*, 17(5), pp. 607–627. doi: 10.1177/1367006912453810.

Rothman, J. (2009) 'Pragmatic deficits with syntactic consequences?: L2 pronominal subjects and the syntax–pragmatics interface', *Journal of Pragmatics*, 41(5), pp. 951–973. doi: 10.1016/j.pragma.2008.07.007.

Rothman, J. and Slabakova, R. (2018) 'The generative approach to SLA and its place in modern second language studies', *Studies in Second Language Acquisition*, 40(2), pp. 417–442. doi: 10.1017/S0272263117000134.

Rumelhart, D. E. and McClelland, J. L. (1982) 'An interactive activation model of context effects in letter perception: II. The contextual enhancement effect and some tests and extensions of the model', *Psychological Review*, 89, pp. 60–94.

Sagarra, N. and Ellis, N. C. (2013) 'From Seeing Adverbs to Seeing Verbal Morphology', *Studies in Second Language Acquisition*, 35(2), pp. 261–290. doi: 10.1017/S0272263112000885.

Sagarra, N. and Herschensohn, J. (2010) 'The role of proficiency and working memory in gender and number agreement processing in L1 and L2 Spanish', *Lingua*, 120(8), pp. 2022–2039. doi: 10.1016/j.lingua.2010.02.004.

Sato, M. and Felser, C. (2007) *Sensitivity to semantic and morphosyntactic violations in L2 sentence processing: Evidence from speeded grammaticality judgments*. University of Essex.

Scherag, A. (2004) 'The effects of late acquisition of L2 and the consequences of immigration on L1 for semantic and morpho-syntactic language aspects', *Cognition*, 93(3), pp. B97–B108. doi: 10.1016/j.cognition.2004.02.003.

Schmid, M. S., Gilbers, S. and Nota, A. (2014) 'Ultimate attainment in late second language acquisition: Phonetic and grammatical challenges in advanced Dutch–English bilingualism', *Second Language Research*, 30(2), pp. 129–157. doi: 10.1177/0267658313505314.

Schustack, M. W., Ehrlich, S. F. and Rayner, K. (1987) 'Local and global sources of contextual facilitation in reading', *Journal of Memory and Language*, 26, pp. 322–340.

Schütze, C. T. and Sprouse, J. (2014) 'Judgment data', in Podesva, R. J. and Sharma, D. (eds) *Research Methods in Linguistics*. Cambridge: Cambridge University Press, pp. 27–50. doi: 10.1017/CBO9781139013734.004.

Schwartz, B. D. and Sprouse, R. A. (1994) 'Word order and nominative case in nonnative language acquisition: a longitudinal study of (L1 Turkish) German interlanguage.', in Hoekstra, T. and Schwartz, B. D. (eds) *Language acquisition studies in generative grammar*. Amsterdam: John Benjamins Publishing Company, pp. 317–368.

Schwartz, B. D. and Sprouse, R. A. (1996) 'L2 cognitive states and the Full Transfer/Full Access model', *Second Language Research*, 12(1), pp. 40–72. doi: 10.1177/026765839601200103.

Seidenberg, M. S. and McClelland, J. L. (1989) 'A distributed, developmental model of word recognition and naming', *Psychological Review*, 96, pp. 523–568.

Sharwood Smith, M. and Truscott, J. (2005) 'Stages or Continua in Second Language Acquisition: A MOGUL Solution', *Applied Linguistics*, 26(2), pp. 219–240. doi: 10.1093/applin/amh049.

Sharwood Smith, M. and Truscott, J. (2014) *The Multilingual Mind: A Modular Processing Perspective*. Cambridge: Cambridge University Press.

Shiffrin, R. M. and Schneider, W. (1977) 'Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory.', *Psychological Review*, 84(2), pp. 127–190. doi: 10.1037/0033-295X.84.2.127.

Shimanskaya, E. and Slabakova, R. (2017) 'Re-assembling objects: a new look at the L2 acquisition of pronominal clitics', *Bilingualism: Language and Cognition*, 20(3), pp. 512–529. doi: 10.1017/S1366728915000784.

Slabakova, R. (2013) 'What is easy and what is hard to acquire in a second language', in García Mayo, M. del P., Juncal Gutiérrez, M., and Martínez-Adrián, M. (eds) *Contemporary Approaches to Second Language Acquisition*. Amsterdam: John Benjamins Publishing Company, pp. 5–28. doi: 10.1075/aals.9.04ch1.

Slabakova, R. (2015) 'The effect of construction frequency and native transfer on second language knowledge of the syntax–discourse interface', *Applied Psycholinguistics*, 36(3), pp. 671–699. doi: 10.1017/S0142716413000386.

Slabakova, R. and Ivanov, I. (2011) 'A more careful look at the syntax-discourse interface', *Lingua*, 121(4), pp. 637–651. doi: 10.1016/j.lingua.2010.05.003.

Slabakova, R., Kempchinsky, P. and Rothman, J. (2012) 'Clitic-doubled left dislocation and focus fronting in L2 Spanish: A case of successful acquisition at the syntax–discourse interface', *Second Language Research*, 28(3), pp. 319–343. doi: 10.1177/0267658312447612.

Slabakova, R., Leal, T. L. and Liskin-Gasparro, J. (2014) 'We Have Moved On: Current

Concepts and Positions in Generative SLA', *Applied Linguistics*, 35(5), pp. 601–606. doi: 10.1093/applin/amu027.

Sorace, A. (2003) 'Near-Nativeness', in Long, C. J. and Doughty, M. H. (eds) *The Handbook of Second Language Acquisition*. Oxford, UK: Blackwell Publishing Ltd, pp. 130–151. doi: 10.1002/9780470756492.ch6.

Sorace, A. (2004) 'Gradience at the Lexicon–Syntax Interface: Evidence from Auxiliary Selection and Implications for Unaccusativity', in Alexiadou, A., Anagnostopoulou, E., and Everaert, M. (eds) *The Unaccusativity Puzzle: Explorations of the Syntax-Lexicon Interface*. Oxford: Oxford University Press, pp. 243–268. doi: 10.1093/acprof:oso/9780199257652.003.0010.

Sorace, A. (2005) 'Syntactic optionality at interfaces', in Cornips, L. and Corrigan, K. (eds) *Syntax and variation: Reconciling the biological and the social*. Amsterdam: John Benjamins Publishing Company, pp. 46–111.

Sorace, A. (2011) 'Pinning down the concept of "interface" in bilingualism', *Linguistic Approaches to Bilingualism*, 1(1), pp. 1–33. doi: 10.1075/lab.1.1.01sor.

Sorace, A. (2012) 'Pinning down the concept of interface in bilingual development', *Linguistic Approaches to Bilingualism*, 2(2), pp. 209–217. doi: 10.1075/lab.2.2.04sor.

Sorace, A. (2014) 'Multiple grammars: Old wine in old bottles', *Second Language Research*, 30(1), pp. 71–74. doi: 10.1177/0267658313517846.

Sorace, A. and Filiaci, F. (2006) 'Anaphora resolution in near-native speakers of Italian', *Second Language Research*, 22(3), pp. 339–368. doi: 10.1191/0267658306sr271oa.

Sorace, A. and Serratrice, L. (2009) 'Internal and external interfaces in bilingual language development: Beyond structural overlap', *International Journal of Bilingualism*, 13(2), pp. 195–210. doi: 10.1177/1367006909339810.

Soutet, O. (2000) *Le Subjonctif*. Paris: Editions Ophrys.

Starke, M. (2010) 'Nanosyntax. A short primer to a new approach to language', *Nordlyd*, 36(1), pp. 1–6. doi: 10.7557/12.213.

Starke, M. (2014) 'Towards elegant parameters: Language variation reduces to the size of lexically-stored trees', in Picallo, M. C. (ed.) *Linguistic Variation in the Minimalist Framework*. Oxford: Oxford University Press, pp. 140–152. doi: 10.1093/acprof:oso/9780198702894.003.0007.

Starr, M. S. and Rayner, K. (2001) 'Eye movements during reading: some current controversies', *Trends in Cognitive Sciences*, 5(4), pp. 156–163. doi: 10.1016/S1364-6613(00)01619-3.

Staub, A. (2015) 'Reading Sentences: Syntactic Parsing and Semantic Interpretation', in Pollatsek, A. and Treiman, R. (eds) *The Oxford Handbook of Reading*. Cary: Oxford University Press, pp. 202–216.

Stepanov, A., Andreetta, S., Stateva, P., Zawiszewski, A. and Laka, I. (2019) 'Anomaly detection in processing of complex syntax by early L2 learners', *Second Language Research*, p. 026765831982706. doi: 10.1177/0267658319827065.

Stewart, A. J., Haigh, M. and Kidd, E. (2009) 'An investigation into the online processing of counterfactual and indicative conditionals', *Quarterly Journal of Experimental Psychology*, 62(11), pp. 2113–2125. doi: 10.1080/17470210902973106.

Stowell, T. (1993) *Syntax of tense*. Los Angeles.

Suzuki, Y. (2017) 'Validity of new measures of implicit knowledge: Distinguishing implicit knowledge from automatized explicit knowledge', *Applied Psycholinguistics*, 38(5), pp. 1229–1261. doi: 10.1017/S014271641700011X.

Svenonius, P. (2019) 'Syntactic Features', in Aronoff, M. (ed.) *Oxford Research Encyclopedia of Linguistics*. Oxford: Oxford University Press, pp. 1–37. doi: 10.1093/acrefore/9780199384655.013.179.

Swets, B., Desmet, T., Clifton, C. and Ferreira, F. (2008) 'Underspecification of syntactic ambiguities: Evidence from self-paced reading', *Memory & Cognition*, 36(1), pp. 201–216. doi: 10.3758/MC.36.1.201.

Taraban, R. and McClelland, J. L. (1988) 'Constituent attachment and thematic role assignment in sentence processing: Influences of content based expectations', *Journal of Memory and Language*, 27(6), pp. 597–632. doi: doi:10.1016/0749- 596x(88)90011-3.

Team, R. D. C. (2014) 'R: A language and environment for statistical computing.' Vienna, Austria: R Foundation for Statistical Computing. Available at: http://www.rproject.org.

Tokowicz, N. and Warren, T. (2010) 'Beginning adult L2 learners' sensitivity to morphosyntactic violations: A self-paced reading study', *European Journal of Cognitive Psychology*, 22(7), pp. 1092–1106. doi: 10.1080/09541440903325178.

Tolentino, L. C. and Tokowicz, N. (2011) 'Across Languages, Space, and Time', *Studies in Second Language Acquisition*, 33(1), pp. 91–125. doi: 10.1017/S0272263110000549.

Traxler, M. J., Pickering, M. J. and Clifton, C. (1998) 'Adjunct Attachment Is Not a Form of Lexical Ambiguity Resolution', *Journal of Memory and Language*, 39(4), pp. 558–592. doi: 10.1006/jmla.1998.2600.

Traxler, M. J. and Tooley, K. M. (2007) 'Lexical mediation and context effects in sentence processing', *Brain Research*, 1146, pp. 59–74. doi: 10.1016/j.brainres.2006.10.010.

Tsimpli, I. M. and Dimitrakopoulou, M. (2007) 'The Interpretability Hypothesis: evidence from wh-interrogatives in second language acquisition', *Second Language Research*, 23(2), pp. 215–242. doi: 10.1177/0267658307076546.

Tsimpli, I. M. and Sorace, A. (2006) 'Differentiating interfaces: L2 performance in syntax-semantics and syntax-discourse phenomena.', in Bamman, D., Magnitskaia, T., and

Zaller, C. (eds) *Proceedings of the 30th annual Boston University Conference on Language Development, BUCLD 30*. Somerville, MA: Cascadilla Press, pp. 653–664.

Ullman, M. T. (2015) 'The declarative/procedural model: A neurobiologically motivated theory of first and second language.', in VanPatten, B. and Williams, J. (eds) *Theories in second language acquisition: An introduction (2nd ed.)*. New York: Routledge, pp. 135–158.

VanAssche, E., Drieghe, D., Duyck, W., Welvaert, M. and Hartsuiker, R. J. (2011) 'The influence of semantic constraints on bilingual word recognition during sentence reading', *Journal of Memory and Language*, 64(1), pp. 88–107. doi: 10.1016/j.jml.2010.08.006.

VanPatten, B. (2007) 'Input Processing in Adult Second Language Acquisition', in *Theories in second language acquisition: An introduction*. Mahwah, NJ: Lawrence Erlbaum Associates., pp. 115–135.

VanPatten, B. (2011) 'Stubborn syntax: How it resists explicit teaching and learning', in Sanz, C. and Leow, R. P. (eds) *Implicit and explicit language learning: Conditions, processes, and knowledge in SLA and bilingualism*. Washington, DC: Georgetown University Press, pp. 9–21.

VanPatten, B. and Jegerski, J. (2010) 'Second language processing and parsing', in VanPatten, B. and Jegerski, J. (eds) *Research in Second Language Processing and Parsing*. Amsterdam: John Benjamins Publishing Company, pp. 3–24. doi: 10.1075/lald.53.01van.

VanPatten, B. and Rothman, J. (2015) 'What does current generative theory have to say about the explicit-implicit debate?', in Rebuschat, P. (ed.) *Implicit and Explicit Learning of Languages*. Amsterdam: John Benjamins Publishing Company, pp. 91–116.

Villegas Erce, A. (2014) *The role of L2 English immersion in the processing of L1 Spanish sentence complement/relative clause ambiguities*. The Pennsylvania State University.

Waller, T. (2017) *The subjunctive in Present-Day English*. University College London.

White, L. (2000) 'Second language acquisition: From initial to final state', in Archibald, J. (ed.) *Second Language Acquisition and Linguistic Theory*. Oxford: Blackwell Publishing Ltd, pp. 130–55.

White, L. (2009) 'Grammatical theory: Interfaces and L2 knowledge', in Ritchie, W. C. and Bhatia, T. K. (eds) *The New Handbook of Second Language Acquisition*. Leeds, UK: Emerald Group Publishing Limited, pp. 49–68.

White, L. (2011) 'Second language acquisition at the interfaces', *Lingua*, 121(4), pp. 577–590. doi: 10.1016/j.lingua.2010.05.005.

White, S. J., Rayner, K. and Liversedge, S. P. (2005) 'The influence of parafoveal word length and contextual constraint on fixation durations and word skipping in reading', *Psychonomic Bulletin and Review*, 12, pp. 466–471.

Whitford, V. and Titone, D. (2019) 'Lexical entrenchment and cross-language activation: Two sides of the same coin for bilingual reading across the adult lifespan', *Bilingualism: Language and Cognition*, 22(1), pp. 58–77. doi: 10.1017/S1366728917000554.

Whong, M., Marsden, H. and Gil, K.-H. (2013) 'How We Can Learn from Acquisition: The Acquisition-Learning Debate Revisited', in *eedings of the 12th Generative Approaches to Second Language Acquisition Conference (GASLA 2013)*. Somerville, MA: Cascadilla Proceedings Project, pp. 203–210.

Williams, J. N. (2006) 'Incremental interpretation in second language sentence processing', *Bilingualism: Language and Cognition*, 9(1), pp. 71–88. doi: 10.1017/S1366728905002385.

Williams, J. N. (2009) 'Implicit learning in second language acquisition', in Ritchie, W. C. and Bhatia, T. K. (eds) *The new handbook of second language acquisition*. Bingley, UK: Emerald Group Publishing Limited, pp. 319–353.

Yang, C. (2004) 'Universal Grammar, statistics or both?', *Trends in Cognitive Sciences*, 8(10), pp. 451–456. doi: 10.1016/j.tics.2004.08.006.

Yang, C. (2018) 'A formalist perspective on language acquisition', *Linguistic Approaches to Bilingualism*, 8(6), pp. 665–706. doi: 10.1075/lab.18014.yan.

Zufferey, S., Mak, W., Degand, L. and Sanders, T. (2015) 'Advanced learners' comprehension of discourse connectives: The role of L1 transfer across on-line and off-line tasks', *Second Language Research*, 31(3), pp. 389–411. doi: 10.1177/0267658315573349.

Zwaan, R. A. and Radvansky, G. A. (1998) 'Situation models in language comprehension and memory.', *Psychological Bulletin*, 123(2), pp. 162–185. doi: 10.1037/0033-2909.123.2.162.