languages

MDPI

Article

L2 Knowledge of the Obligatory French Subjunctive: Offline Measures and Eye Tracking Compared

Amber Dudley [1,*] and Roumyana Slabakova [1,2]

[1] Department of Modern Languages and Linguistics, University of Southampton, Southampton SO17 1BJ, UK; R.Slabakova@soton.ac.uk
[2] NTNU the Norwegian University of Science and Technology, 7491 Trondheim, Norway
* Correspondence: A.Dudley@southampton.ac.uk

Abstract: Extensive research has shown that second language (L2) learners find it difficult to apply grammatical knowledge during real-time processing, especially when differences exist between the first (L1) and L2. The current study examines the extent to which British English-speaking learners of French can apply their grammatical knowledge of the French subjunctive during real-time processing, and whether this ability is modulated by the properties of the L1 grammar, and/or proficiency. Data from an acceptability judgment task and an eye-tracking during reading experiment revealed that L2 learners had knowledge of the subjunctive, but were unable to apply this knowledge when reading for comprehension. Such findings therefore suggest that L2 knowledge of the subjunctive, at least at the proficiency levels tested in this study, is largely metalinguistic (explicit) in nature and that reduced lexical access and/or limited computational resources (e.g., working memory) prevented learners from fully utilising their grammatical representations during real-time processing.

Keywords: second language acquisition; second language processing; first language influence; eye movement data; subjunctive

Citation: Dudley, Amber, and Roumyana Slabakova. 2021. L2 Knowledge of the Obligatory French Subjunctive: Offline Measures and Eye Tracking Compared. *Languages* 6: 31. https://doi.org/10.3390/languages6010031

Received: 31 December 2020
Accepted: 8 February 2021
Published: 17 February 2021

1. Introduction

The dynamic nature of morphosyntactic variation in both comprehension and production has been well documented in the field of second language (L2) acquisition and processing (Franceschina 2005; Grüter et al. 2012; Hopp 2010; Keating 2009; Lardiere 1998, 2009; Slabakova 2018; Keating 2010). Over the past two decades and in response to a substantial and ever-expanding body of experimental work, the field as a whole has proposed and subsequently refined numerous generatively grounded theories of second language acquisition. Specifically, these theories have sought to elucidate the deterministic factors that contribute to L1–L2 differences in ultimate attainment that access to Universal Grammar (UG) alone cannot explain. Of particular note and interest to the current study is the Feature Reassembly Hypothesis (Lardiere 2009) and the subsequent studies testing its predictions (see, e.g., Cho and Slabakova 2014; Domínguez et al. 2017; Gil and Marsden 2013; Guijarro-Fuentes 2012; Hwang and Lardiere 2013). The Feature Reassembly Hypothesis (FRH) predicts that the initial state of L2 learners' interlanguage grammar consists of "an entrenched system of morphosyntactic features already assembled into lexical items" (Hwang and Lardiere 2013, p. 58). This consequently requires L2 learners to compare and contrast L1 and L2 properties, interpreted as formal featural[1] configurations, in order to map the L1 morphosyntactic form(s) onto the L2 counterpart(s). In the (likely) event of form-to-meaning mapping mismatches, the L2 learner is required to reconfigure existing L1 featural compositions into new L2 featural bundles, along with the appropriate lexical items. The most problematic linguistic structures for L2 learners are those with

[1] This paper interprets formal (i.e., functional) features as the "formal properties of syntactic objects which determine how they behave with respect to syntactic constraints and operations" (Svenonius 2019, p. 1).

complex and highly divergent L1–L2 featural compositions. In other words, the greater the difference between the L1 and the L2, the more challenging the acquisition task will be.

While this body of research has undeniably enriched our understanding of how the mind represents and uses a second language, particularly in terms of the role of L1 influence, most approaches have typically relied on offline measures, such as grammaticality judgment tasks, in order to probe L2 knowledge. As such, few studies have adopted online measures, such as eye-tracking during reading. Consequently, we as a field still have a lot more to learn about how L1 grammatical knowledge, particularly L1–L2 differences in feature configurations, impacts on L2 representations during real-time processing. Furthermore, it is not certain to what extent judgment tasks really do tap into L2 learners' underlying linguistic representations, with some scholars suggesting that such tasks are the by-product of metalinguistic (i.e., taught) knowledge about language (Suzuki 2017). This is particularly problematic in the context of structures that are frequently and explicitly taught in the L2 classroom, such as the subjunctive. However, online methods, such as eye-tracking, have been argued to measure real-time interpretation and thus have the potential to be a more accurate measure of underlying implicit knowledge, depending on the experimental design used (Keating and Jegerski 2015). For example, when participants are required to simply read sentences for comprehension, without having to press a button to see the next part of the sentence, or assess the grammaticality of the sentence in question, the amount of time available to access explicitly learned, metalinguistic knowledge is reduced, which in turn requires participants to rely more on their implicit knowledge.

To what extent, then, can online methods, such as eye-tracking during reading, measure L2 representations? One could argue that these methods are simply a measure of processing ability. However, if we follow the logic of Fodor (1998a, 1998b) Parsing to Learn Hypothesis and its recent application to L2 acquisition by Dekydtspotter and Renaud (2014), these methods have the potential to weigh in on existing debates not only about L2 processing but also about L2 representations. Specifically, Dekydtspotter and Renaud's application of Fodor's Parsing to Learn Hypothesis assumes that L2 processing patterns are determined not by language-specific parsing routines, but rather by the grammar sub-serving the parser. Put simply, when a comprehender, regardless of their L1, has access to a particular featural configuration in the grammar/mental lexicon, they will be able to accurately parse the linguistic structure in question and the input, more generally. When such a featural configuration is not present in the grammar, then parsing failure will inevitably arise. This will in turn trigger feature reassembly and subsequent acquisition. Given research on the eye–mind link (Rayner 2009; Liversedge and Findlay 2000), such an argument suggests that the influence of the L1 morphosyntax should be reflected in the eye movement patterns of L2 learners. We are still mindful, however, of the fact that unexpected processing patterns may arise even if target-like featural configurations are in place due to delayed lexical access and/or limited computational/cognitive resources, as argued by limited capacity models (e.g., McDonald 2006; Hopp 2010; Sorace 2011; Dekydtspotter and Renaud 2014). With this in mind, the current study combines both judgment (offline) and eye movement (online) data to investigate whether L2 learners are able to apply their offline knowledge of the subjunctive during online processing. It must be noted at this point that there is extensive debate in the field of L2 acquisition about the extent to which offline tasks, such as judgment tasks, reflect learners' implicit knowledge of language, with factors such as whether the task is timed or untimed, and the nature of L2 learners' exposure (i.e., instructed vs. naturalistic), playing an important role (cf., Ionin and Zyzik 2014; Plonsky et al. 2020).

The current article therefore adopts a processing-based extension of the Feature Reassembly Hypothesis in order to investigate the extent to which L1–L2 differences in feature configurations modulate the L2 representations and processing of the subjunctive of British English-speaking second language learners of French, using a combination of acceptability judgment and eye-tracking during reading data. Exploring the L2 acquisition of the French subjunctive by L1 British speakers is a pertinent test case, given the morphosyntactic differences between English and French, which we will discuss in the following section. Furthermore, although many studies have explored the various factors contributing to this developmental delay, including the deterministic role of proficiency (Ayoun 2013; McManus and Mitchell 2015), licensing conditions/trigger type (Ayoun 2013; McManus and Mitchell 2015; Bartning 2008; Howard 2008), and the role of learning context (McManus and Mitchell 2015; Howard 2008), very few have directly considered the role of L1 influence. Although Ayoun (2013) did not directly explore L1 influence, she did investigate L2 learners' knowledge of the subjunctive according to the lexical–semantic properties of the matrix predicate and found that L2 learners performed most accurately with order/interdiction (i.e., directive) predicates in comparison to judgment, emotion, wish/regret, and doubt/impossibility predicates, despite the relative infrequency of directive predicates in the input (see, e.g., Kastronic 2016). We therefore believe that such a finding merits further systematic investigation.

2. Morphosyntactic Differences in Subjunctive Use between British English and French

If we are to describe the acquisition task within a Feature Reassembly framework in detail, we must first engage in a cross-linguistic comparison of subjunctive licensing conditions in both English and French, focusing in particular on the featural configurations of the relevant structures.

2.1. Subjunctive in French

The current article focuses on the use of the subjunctive[2] in complementiser phrases where mood choice results from the lexical–semantic properties of the matrix predicate. In the literature, this type of subjunctive is broadly referred to as the volitional subjunctive (Quer 2009; Giannakidou 1998, 2009). Henceforth, we refer to it as the obligatory subjunctive.

In French, verbs of saying (e.g., dire "say", observer "observe") and epistemic predicates (e.g., se render compte "realise", se rappeler "remember", penser "think") select indicative complementiser phrases (CPs), whereas emotive–factive (e.g., regretter "regret"), desire (e.g., préférer "prefer", souhaiter "wish", vouloir "want") and directive (e.g., demander "ask", exiger "demand") predicates take subjunctive CPs (Baunaz 2017). For example, in sentence (1) below, the matrix verb, penser "to think", requires an indicative complementiser phrase, but the matrix verb, vouloir "to want", in sentence (2), requires a subjunctive complementiser phrase.

(1) 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."

(2) 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."

Over the years, many proposals have been advanced in order to account for mood alternation in the complementiser phrases of subjunctive-selecting predicates. According to Baunaz (2017), all predicates selecting tensed CPs carry a set of semantic features, and it is these semantic features which trigger mood selection in French complementiser phrases (CPs). The semantic decomposition of these verbs is shown in Table 1 below. As can be seen from Table 1, all subjunctive-selecting predicates carry an emotive feature. Baunaz (2017) therefore proposes that it is the presence of this very feature that determines mood selection in French.

[2] Following Portner (2018, p. 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".

Table 1. Semantic decomposition of verbs selecting tensed CPs in French.

	Saying	Epistemic	Emotive–Factive	Desire	Directive
Mood	ind.	ind.	subj.	subj.	subj.
Sentient[3]	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).

Quer (1998, 2001, 2009), on the other hand, argues that mood selection results from a shift in the model evaluating the truth-value of the proposition in the embedded clause. In the case of the obligatory subjunctive, the model shift occurs as a result of the lexical–semantic properties of the matrix predicate. Specifically, the model shifts from the default model (i.e., the actual world as described by the speaker's epistemic model) to the bouletic model of the matrix subject. This means that the shift occurs not only with respect to the model's anchor (i.e., from the speaker to the matrix subject), but in the type of model (i.e., from an extensional epistemic model to a strongly intensional bouletic model).[4]

Although many consider the subjunctive to be a largely semantic phenomenon, Kempchinsky (2009) suggests that subjunctive complementiser phrases represent a purely syntactic phenomenon, incorporating the following three functional projections: the Fin and Force heads in the left periphery and a Mood head in the inflectional field. According to Kempchinsky (2009), the model shift, which we have just described, is represented syntactically via an uninterpretable modal feature, [M], in the Force head of the left periphery. This uninterpretable feature is then adopted by FinP, a projection which is argued to carry information about the discourse context (Bianchi 2001). This is because the deixis centre of the subjunctive is located within the bouletic intensional model and therefore "only indirectly linked to the extensional speech context" (Iverson et al. 2008, p. 138). In line with the minimalist theory of feature checking, the uninterpretable features in both FinP and ForceP are checked and deleted from the derivation by means of an interpretable feature in the head of the morphosyntactic projection, MoodP. Note that it is the latter interpretable feature that gives rise to the verbal inflections for mood. Such an analysis can be represented as in Figure 1.

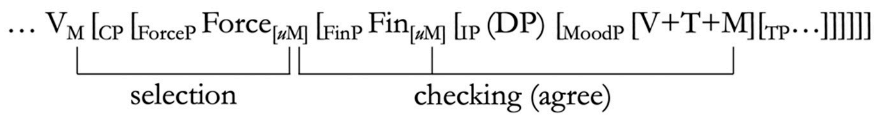

Figure 1. Syntactic analysis adapted from Iverson et al. (2008, p. 139).

While this analysis considers the subjunctive to be a purely syntactic phenomenon, it does not neglect the semantic contribution. Specifically, Kempchinsky argues that the semantic dimension of the subjunctive results from the lexical–semantic properties of the selecting predicate, and this is precisely what Baunaz (2017) describes. The current study therefore adopts Baunaz (2017) approach to compositionality to describe the semantic contribution of the subjunctive and combines it with Kempchinsky's syntactic analysis of the subjunctive.

It must be noted that while the above analyses treat the subjunctive as a morphosyntactic opposition, where one mood excludes the other, many have argued (e.g., Gudmestad and Edmonds 2015 for Hexagonal French and Poplack et al. 2013 for Canadian) that the subjunctive is an example of a variable structure.

3 Baunaz (2017, p. 19) uses the term, sentience, to refer to "a cognitive state, emotion or perception".

4 A strongly intensional bouletic model represents a set of worlds where the desires of the matrix subject are expressed (von Fintel 2006, p. 2).

2.2. Subjunctive in English

The semantic contribution of the subjunctive is particularly relevant when you consider the cross-linguistic differences between English and French in terms of the licensing conditions of the subjunctive. Unlike French, where the subjunctive is used in the complementiser phrases of desire, directive, and emotive–factive predicates, the licensing conditions for the subjunctive are much more restrictive in English, particularly in British English.[5] For example, English only allows subjunctive complementiser phrases with directive predicates, as in sentences (3)–(5) below, but not with desire and emotive–factive predicates, and even then, this is in alternation with indicative complementiser phrases. In other words, both the indicative and the subjunctive are equally acceptable in the complementiser phrases of English directive predicates.[6] Nevertheless, the fact remains that British English speakers, in large part, have access in their grammar(s)[7] to the featural specifications for the subjunctive in the context of directive predicates.

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

It is important to highlight here that while desire predicates do not mark modality morphosyntactically, this does not mean that their complementiser phrases do not carry a modal feature. For example, in English, desire predicates are typically followed by infinitival complementiser phrases, as in sentences (6)–(7) below. Note, however, that the complementiser, for, is often omitted, as in (8). This does not mean that it is not present in the syntax; instead, it is assumed to be phonologically null (i.e., not pronounced), given its presence in clefted sentences, such as (9).

(6) I really prefer for Lewis to be the main presenter.
(7) What I suggest is for Lewis to be the main presenter.
(8) The professors want [$_{CP}$ [$_{C}$ Ø] [James to present the paper]]
(9) What the professors want is for James to present the paper.
(10) *What the professors want is James to present the paper.

Iverson et al. (2008) have, therefore, argued that the infinitival complementiser, for, should be analysed as a marker of intensional modality, whereby it joins the syntax as the head of FinP in the left periphery. Such an analysis is schematised in Figure 2 below.

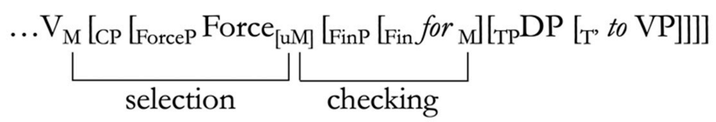

Figure 2. The syntactic structure of for–to infinitival phrases in English.

Such an analysis therefore suggests that for serves as a morphosyntactic feature, which is able to detect the uninterpretable modal feature in Force. Nevertheless, we must highlight at this point that this does indeed contradict the somewhat narrow definition of mood as inflectional morphology.

Finally, while directive and desire CPs express modality in one form or another, emotive–factive predicates categorically do not express modality and, as we will discuss in the next section, this has important implications for the acquisition task.

---

[5] See Dudley (2020) for an empirical study investigating the licensing contexts for the subjunctive in British English.

[6] It could be argued that the acceptance of both indicative and subjunctive complementiser phrases reflects the presence of two competing grammars in the minds of British English, with indicative complements forming part of the informal grammar and subjunctive complements part of the formal grammar and with access to each grammar depending on sociolinguistic/contextual factors.

[7] It is indeed possible that delayed L1 acquisition (as is the case of the subjunctive in Romance languages) may result in the development of more than one grammar (see, e.g., Kato et al. 2009; Pires and Rothman 2009; Zribi-Hertz 2011).

2.3. Acquisition Task Predictions

As we discussed in the introduction, the current article seeks to adopt a processing-based extension of the Feature Reassembly Hypothesis. To recap, the Feature Reassembly Hypothesis predicts that the initial state of L2 learners' interlanguage grammar consists of "an entrenched system of morphosyntactic features already assembled into lexical items" (Hwang and Lardiere 2013, p. 58). This subsequently requires L2 learners to compare and contrast L1 and L2 properties, interpreted as formal featural configurations in the context of the FRH, in order to map L1 morphosyntactic form(s) onto the L2 counterpart(s). In the (likely) event of form-to-meaning mapping mismatches, the L2 learner is required to reconfigure existing L1 featural compositions into new L2 featural bundles, along with the appropriate lexical items. The most problematic linguistic structures for L2 learners are therefore said to be those with complex and divergent L1–L2 featural compositions. In other words, the greater the difference between the L1 and the L2, the more challenging the acquisition task will be.

As we have just established in the previous subsection, there are substantial differences in subjunctive use between English and French, on both a morphosyntactic and semantic level. Nevertheless, the core assumption underlying the acquisition task predictions presented here is that British English-speaking L2 learners of French already have access to the functional category, MoodP, and the relevant modal features via directive predicates, as we have described previously. As a result, L2 learners must recognise that the L2 French maps these features onto lexical items differently from the L1 British English. In this section, we present the acquisition task on a predicate-by-predicate basis for clarity of expression. However, a visual summary of the acquisition task is presented in Figure 3.

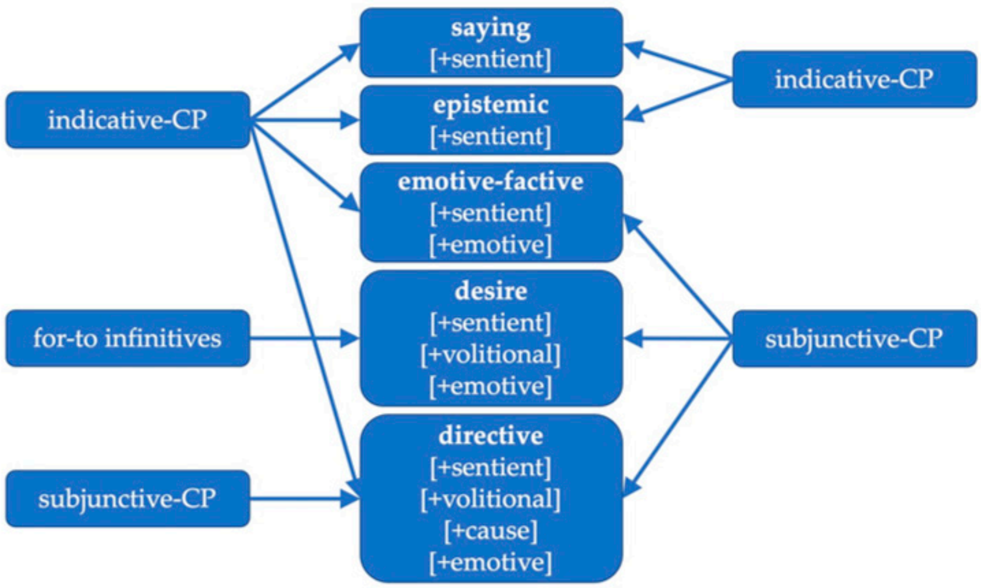

Figure 3. Mapping of the semantic features of verbs selecting tensed CPs (middle column) onto corresponding forms in English (left column) and French (right column).

As we know, both English and French license subjunctive CPs with directive predicates. Cross-linguistic differences arise, however, as a result of the optionality in English. That is, both indicative and subjunctive CPs are equally grammatical in English, dependent on register, of course. What the learner must do, therefore, is not allow a superset of two possible CPs as in the L1, but a subset of only subjunctive CPs in the L2.

With desire predicates, it is assumed that modality is marked both in English and French, albeit onto different surface forms. That is, where English marks modality via for-to infinitival CPs, French marks modality via subjunctive CPs. L2 learners must therefore reassemble the interpretable modal feature, [M], in FinP, mapped onto the complementiser, for, in the L1 to the inflectional field of MoodP, while adding a T(ense) feature. Alternatively, one could conceive of the learning task as extending the licensing contexts of the featural compositions, currently mapped onto directive CPs to desire and emotive–factive CPs.

What does this all mean in terms of concrete predictions, then? We predict that Anglophone learners of L2 French will initially treat the indicative as grammatical in all CPs, regardless of the lexical–semantic properties of the matrix predicate, given that this is the closest morpholexical equivalent in their native language. It is possible that some optionality will be observed between the indicative and subjunctive with directive predicates due to L1 influence, but that the indicative will more often than not be preferred, given the relative infrequency of the subjunctive in British English. For acquisition to proceed and for L2 learners to unlearn that the indicative is ungrammatical with these predicates, L2 learners must be exposed to sufficient evidence of the target structure (i.e., the subjunctive) in the input.

As L2 learners become more proficient, we anticipate that target-like sensitivity to subjunctive CPs will first surface with directive predicates (e.g., at intermediate to upper-intermediate levels). As L2 learners gradually remap and reconfigure the featural configurations onto new surface morphology, this will extend first to desire and then emotive predicates (e.g., at upper-intermediate to advanced levels). As we have already discussed, this prediction is based on the idea that L2 learners already mark modality with desire predicates in the L1, albeit via a (null or overt) complementiser rather than inflectional morphology. This is not, however, the case with emotive–factive predicates.

It is essential to highlight at this point that any acquisition fundamentally depends on the availability of unambiguous input. For example, Slabakova (2015, p. 680) argues that input is crucial for feature acquisition and, as such, constructions which are either not salient or infrequent in the input are likely to generate significant acquisition difficulties and potentially even developmental delays. In the case of the subjunctive, the sufficient quality (i.e., whether there are enough examples of the linguistic structure for form-to-meaning mapping to occur) and quantity of input is far from guaranteed and is likely to be a complicating factor (O'Connor DiVito 1997; Poplack et al. 2013).

It is equally possible that other important factors, such as frequency (i.e., L2 learners typically acquire the highly frequent structure, il faut que "it is necessary that", earlier than other subjunctive triggers) and the nature of explicit instruction (i.e., the subjunctive is often taught in relation to categories such as "emotion" and "desire" rather than "directive", for example).

3. Research Questions and Predictions

The current study set out to explore the extent to which the L1 morphosyntax influences the acquisition and processing of mood–modality mismatches in L1 French and L2 French at upper-intermediate to advanced levels of proficiency. Specifically, it sought to address the following research questions and subsequent predictions:

1. Are British English-speaking L2 learners of French sensitive to the syntactic and semantic licensing conditions of the subjunctive in obligatory contexts, in both online and offline comprehension?
2. To what extent is this sensitivity modulated by:
    a. L1–L2 differences in featural configurations and
    b. proficiency?

In terms of the L1 speakers' behaviour, we expected to find significant differences between grammatical (i.e., subjunctive CPs) and ungrammatical (i.e., indicative CPs) sentences, regardless of the lexical–semantic properties of the matrix (i.e., selecting, predicate), both in terms of judgment ratings and reading measures. With judgment ratings, we predicted that grammatical (i.e., subjunctive) sentences would consistently attract higher acceptability ratings than ungrammatical (i.e., indicative) sentences. With reading measures, we anticipated that ungrammatical sentences would consistently attract longer reading times than grammatical sentences, based on previous research by Rayner and colleagues (2004) demonstrating that morphosyntactic violations can have a direct impact on reading measures—most notably, gaze duration.

In terms of the L2 speakers, we expected their behaviour to be less clear cut but, most importantly, modulated by the L1 properties. First and foremost, we anticipated there to be a distinct asymmetry between offline (i.e., judgment) and online (i.e., eye movement) data, in light of previous research (cf., Papadopoulou and Clahsen 2003; Marinis et al. 2005; Shimanskaya and Slabakova 2017; Stepanov et al. 2019). Although we have argued that online knowledge may be modulated by underlying linguistic knowledge, it is likely that delayed lexical access and/or limited computational resources may prevent learners from applying this knowledge during real-time processing. Such an argument is espoused most notably by limited capacity models (e.g., McDonald 2006; Hopp 2010; Sorace 2011; Dekydtspotter and Renaud 2014). Second, we expected that L1–L2 morphosyntactic differences would modulate L2 processing patterns, and, to a lesser extent, judgments, of mood–modality mismatches. Specifically, at the lowest proficiency levels, we predicted that L2 learners would treat indicative CPs as grammatical, with shorter reading times and higher ratings for the ungrammatical (i.e., indicative) condition, as compared to the grammatical (i.e., subjunctive) condition. A slight caveat here is that it is likely that there would be no significant difference between the ungrammatical and grammatical conditions with directive predicates, given the previously predicted optionality between the two moods. Nevertheless, we expected that as L2 learners became more proficient, their processing patterns and judgments would approximate, or perhaps even align with, the L2 norms, first with directive predicates and then desire and emotive–factive predicates, in this specific order. We must note here that target-like sensitivity to mood–modality mismatches is most likely to be found in the spillover rather than the critical region, given the widely reported delayed sensitivity to morphosyntactic anomalies among the L2 learners (see, e.g., Felser and Cunnings 2012; Felser et al. 2012; Lim and Christianson 2015; Boxell and Felser 2017). The predictions are summarised below:

Hypothesis 1 (H1). *L2 learners should be more sensitive to mood–modality mismatches in the offline data than the online data.*

Hypothesis 2 (H2). *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 (H3). *As proficiency increases, L2 learners should exhibit sensitivity to mood–modality mismatches, not only with directive predicates, but also with desire and emotive–factive predicates.*

4. Method

4.1. 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 authors' university. They were paid GBP 10.00 for their participation in the study. Any participants 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 (Mulder and Hulstijn 2011; Pakulak and Neville 2010), 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.

4.1.1. Native Speakers

Data were collected from native speakers of French as a baseline measure of performance against which to evaluate learner performance on the experimental tasks. The participants in the control group (n = 30, 24 females, mean age = 21.82 years, SD = 1.76 years) were mostly English–French bilinguals, whose first exposure to the English language was at approximately 8.67 years of age. At the time of testing, participants had been resident in the United Kingdom for an average of 7.73 months. In a bid to avoid the comparative fallacy (Bley-Vroman 1983), we made the conscious decision to recruit bilingual speakers, and not monolinguals. Previous studies have shown that the parallel activation of two languages in the mind leads to increased competition, which places cognitive demands on the speaker during real-time processing, due to the need to suppress the language not in use (e.g., Green and Abutalebi 2013; Kroll and Bialystok 2013; Baum and Titone 2014). As such, we wanted the control group and L2 learners to be as comparable as possible, the only difference being whether participants were tested in their L1 or L2.

4.1.2. Second Language Learners

We recruited 45 native English-speaking learners of French (34 females, mean age = 21.42, SD = 1.12 years). On average, L2 participants reported first exposure to French at 9.47 years (SD = 2.41 years), having already acquired their first language. At the time of testing, the L2 participants were either in their second year or final year of an undergraduate degree, or a one-year postgraduate degree, studying one or more modern foreign language at the authors' university. This meant that many participants reported knowledge of at least another language, in addition to English and French.[8]

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). It is a standardised test of vocabulary knowledge aimed at learners of upper-intermediate to advanced proficiency.[9] The task in question consisted of an un-speeded 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 then required to identify which of the 84 words were real French words. Based on this task, participants were classed as upper-intermediate to advanced L2 learners of French. Figure 4 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 ($F(1,76) = 257.1$, $p > 0.001$).

Prior to testing, all 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. See Tables A1 and A2 in Appendix A for a summary of these data.

[8] It is likely that knowledge of the subjunctive in another Romance language, such as Spanish, may have positively influenced the knowledge of the subjunctive in (L3 or L_n) French. Since we did not test their knowledge of the subjunctive in additional languages, we were unable to test this experimentally. Future studies, however, should focus on the extent to which knowledge of the subjunctive in another Romance language influences the L2/L_n development of the French subjunctive.

[9] As an anonymous reviewer highlighted, this type of proficiency test raises an important question about whether it is appropriate to use a measure of lexical knowledge as an index of proficiency, particularly as part of a study assessing morphosyntactic knowledge. However, Lemhöfer and Broersma (2012) found that LexTALE scores correlate highly ($rs > 0.6$) with standardised measures of proficiency, such as the Quick Placement Test (QPT), which assesses both morphosyntactic and lexical knowledge.

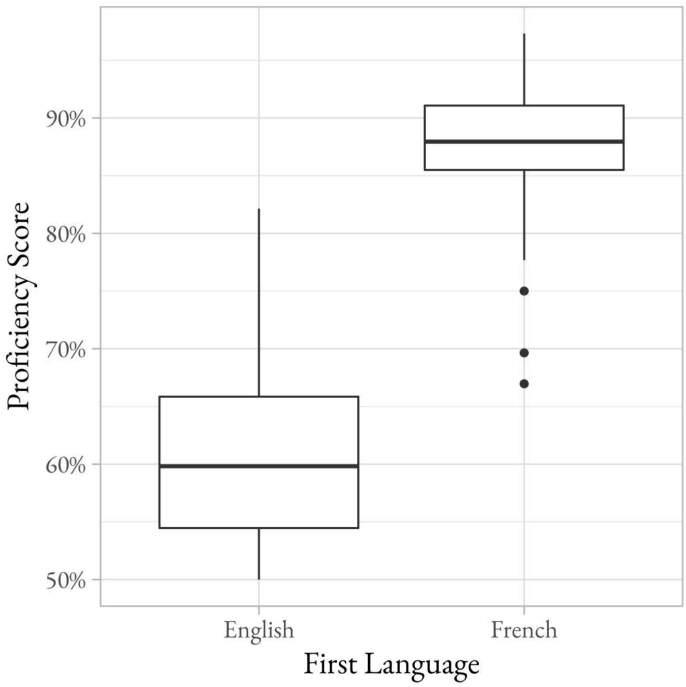

Figure 4. Raw proficiency scores.

4.2. 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 minimised using a chin and forehead rest.

4.3. Materials

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")[10] 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 manipulated was grammaticality (i.e., indicative and subjunctive) in the embedded clause. In each case, only the subjunctive form was grammatical. Examples of test items are provided in Table 2.

The materials included two versions of 36 experimental items: one with a subjunctive CP and the other with an indicative CP. The 72 experimental sentences resulting from the 36 items were counterbalanced across two presentation lists. Each list contained 36 sentences, within which one version of each item was included.

[10] Given the lack of morphological transparency between many indicative and subjunctive forms (O'Connor DiVito 1997; Poplack 1989, 1990; Soutet 2000; Poplack et al. 2013), we specifically chose to select these verbs to mark the indicative–subjunctive distinction.

Table 2. Example test items.

Semantic Properties	Mood	Example
Desire	Indicative	* Cécilia veut que son amie **a** plus de confiance en elle.
	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.
	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.
	Subjunctive	Pierre craint que Sophie **soit** toujours amoureuse de Claude. "Pierre is afraid that Sophie is still in love with Claude."

* Denotes an ungrammatical sentence.

4.4. Procedure

Before each testing session, both L1 and L2 participants filled out a background information questionnaire, the data of which are summarised in Tables A1 and A2 in Appendix A. During the testing session, participants completed three tasks in the following order: an eye-tracking during reading experiment, an acceptability judgment task, and a (lexical) proficiency test.

During both the eye-tracking during reading experiment and the acceptability judgment task, participants read 36 experimental sentences and 36 fillers in each session, which were presented in an individually randomised order. These test items were identical across both the eye-tracking during reading experiment and acceptability judgment task.[11] In the eye-tracking during reading experiment, participants were instructed to silently read each sentence for comprehension at their normal reading rate (for no longer than 10 s) and then answer a comprehension question (within 15 s) on 50% of trials. Comprehension questions focused on the content of the sentence, not the form, and required a yes–no push button response.

Each testing session started with a nine-point calibration procedure, followed by six practice trials. 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 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. Each session lasted approximately 10 minutes.

In the judgment task following the eye-tracking experiment, participants read 36 experimental and 36 filler sentences and were asked to assess the grammaticality of each sentence, using a 7-point scale from 1 to 7, with "1" signifying completely unacceptable and "7" completely acceptable (within 25 s).

4.5. Data Preparation and Analysis

Eye movements were analysed at three regions: the matrix predicate, the critical region, and the spillover region. The critical region consisted of the embedded verb in the complementiser phrase and the spillover region consisted of one to three words following the critical region.[12] The decision to analyse reading times at the spillover region was based on previous studies showing that sensitivity effects are often delayed among non-native populations (Felser and Cunnings 2012; Felser et al. 2012; Lim and Christianson 2015; Boxell and Felser 2017). See Table 3 below for a visualisation of these regions.

11 We acknowledge that this will have undeniably led to repetition effects. We argue, however, that the judgment task was designed to probe L2 learners' metalinguistic knowledge of the grammar.

12 Where possible, we tried to minimise the difference in the length (in characters) of the spillover region between items, which meant that for some items, the spillover region was a single word, whereas for other items, it was a series of (relatively) short words.

Table 3. Regions of interest in the current study.

		Matrix Predicate		Pre-Critical	Critical	Spillover	
subjunctive	Victoria	préfère	que	Claude	soit	responsable	du projet
indicative	Victoria	préfère	que	Claude	est	responsable	du projet

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, fixations that were more than three standard deviations from the mean of each dependent variable were deleted.

To correct the skewed distribution and thus approximate normal distribution, fixation data were log-transformed. Mixed-effects linear regression models were then conducted using the lmer function of the lme4 package (Bates et al. 2015) in the R environment (R Development Core Team 2014). 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 exceeding $\pm$ 1.96 were analysed as significant, and p values were calculated using the lmerTest package (Kuznetsova et al. 2017) in the R environment (R Development Core Team 2014).

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 grammaticality (i.e., mood: indicative (ungrammatical) and subjunctive (grammatical)), matrix semantic property (desire, directive, and emotive–factive), and their interaction. Grammaticality and matrix semantic property were inputted as categorical fixed effects and sum contrasts were coded. For grammaticality, we explored the difference between ungrammatical (indicative) and grammatical (subjunctive) sentences (ungrammatical -1, grammatical 1). For the 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, emotive–factive 1). Directive sentences were chosen as the baseline since these are the only verbs in English that allow the subjunctive mood.

To control for frequency confounds in the critical region, zipf frequency (centred) was inputted as a continuous fixed effect, but only in the critical region models. Trial index was also entered as a continuous fixed factor to control for any confounds relating to the position of each test item within the experiment. 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 in the L2 models, given that most native speakers performed close to ceiling in the proficiency test.

Finally, any significant fixed effects and/or interactions were interpreted using 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.

5. Results

5.1. Judgment

It has been previously stated that the acquisition of the relevant grammatical knowledge "does not automatically entail the processing of something related" (VanPatten and Jegerski 2010, p. 8). Before analysing L2 processing patterns, it was important to first establish that L2 learners had already acquired the relevant grammatical knowledge. We did this by running mixed-effects models on the L1 and L2 judgment data. The predictor variables included in the mixed-effects models were identical to the ones described in Section 4.4. The raw means (and standard deviations) are presented in Table 4 and the estimates from the mixed-effects models for the L1 and L2 speakers in Tables A3 and A4 of Appendix A, respectively.

Table 4. Raw means (standard deviations) for judgment ratings (scale 1–7).

		L1	L2
Desire	Indicative	2.72 (1.80)	4.24 (2.12)
	Subjunctive	6.74 (0.69)	6.30 (1.10)
Directive	Indicative	2.88 (1.94)	4.79 (1.94)
	Subjunctive	6.77 (0.71)	6.19 (1.30)
Emotive–Factive	Indicative	2.77 (1.81)	4.66 (2.08)
	Subjunctive	6.73 (0.81)	6.29 (1.20)

The mixed-effects model revealed a significant main effect of grammaticality among both the L1 speakers (β = 1.94, SE = 0.03, t = 58.48, $p < 0.001$) and the L2 speakers (β = 1.23, SE = 0.06, t = 19.44, $p < 0.001$), with subjunctive sentences consistently being rated as more acceptable than indicative sentences, as expected. Among the L2 speakers, we additionally found a significant interaction between grammaticality and proficiency (β = 0.53, SE = 0.07, t = 7.29, $p < 0.001$), suggesting that their judgments became more categorical as a function of proficiency (see Figure 5).

5.2. Eye Movement Data

Seven eye movement measures were calculated at the matrix predicate, the critical region (i.e., the embedded mood-marked verb), and the spillover region. First fixation duration (FFD) is the duration of the first fixation on a region of interest; gaze duration (GD) is the sum of all fixations on a region of interest before the eyes progress to another word; go-past time (GPT) is the sum of all fixations from the first time the eyes enter the region until leaving it to the right; total reading time (TRT) is the sum of all fixations on a region of interest; the skipping rate (SKIP) is the probability that a reader will skip the region of interest on first pass read; regression-in probability (IN) is the percentage of times a reader regressed back to earlier parts of the sentence from the current region of interest, and regression-out probability (OUT) is the percentage of times a reader regressed back to the current region of interest. Although the current study will report the raw means for skipping rates, it will not present inferential analyses on these, in light of research (e.g., Brysbaert et al. 2005) showing that word length is the strongest influence by far on skipping rates and, as a result, even a difference of two letters can have a significant impact.

The current subsection is structured as follows. We first start by discussing the L1 eye movement data, focusing first on reading measures at the matrix predicate, the critical region, and finally the spillover region. This is then repeated for the L2 eye movement data.

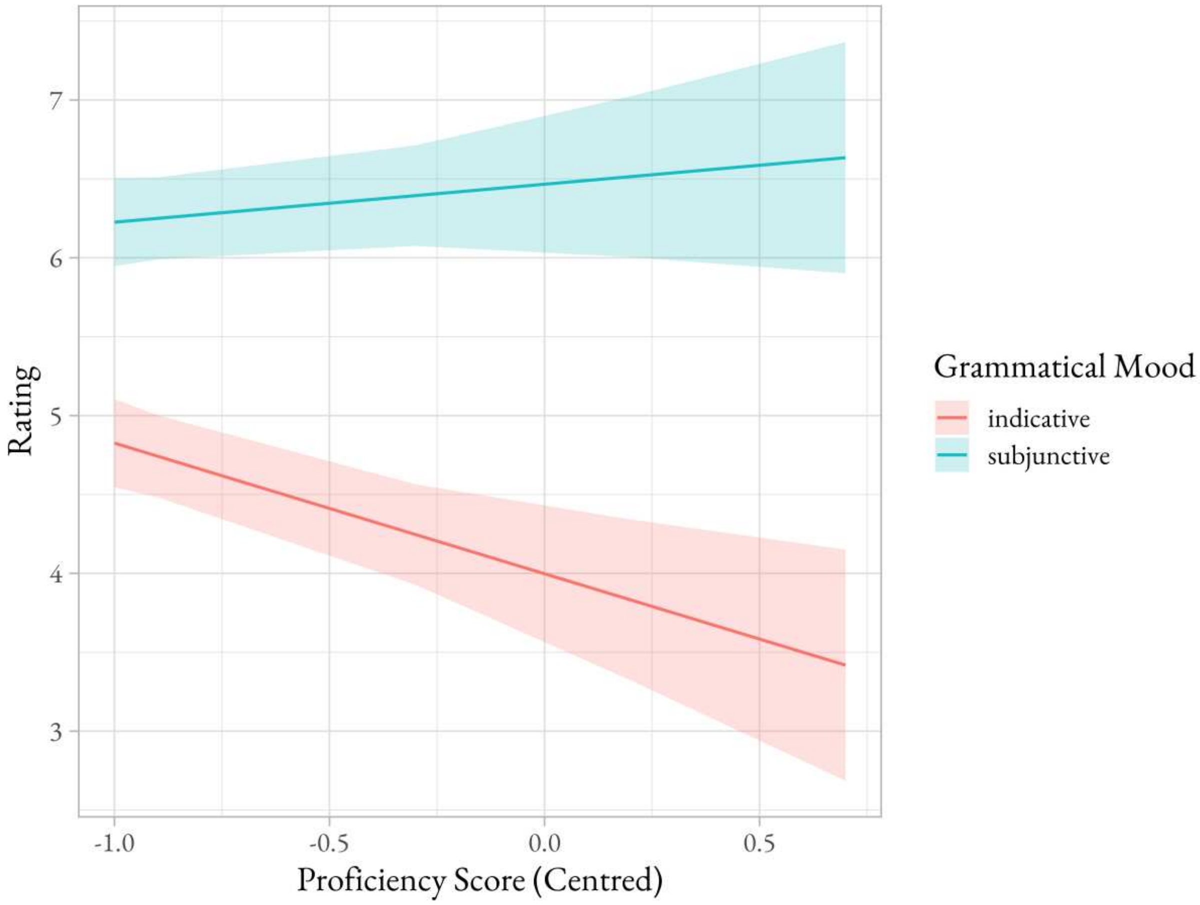

Figure 5. L2 judgment ratings of grammatical (subjunctive) and ungrammatical (indicative) sentences, as a function of proficiency.

5.2.1. L1 Speakers

The raw means and the mixed-effects models for the reading measures are presented in Table 5 below and Table A5 in Appendix A, respectively.

Table 5. Reading time measures for L1 speakers.

			FFD	GD	GOPAST	TRT	SKIP	IN	OUT
Matrix Predicate	Desire	Indicative	214.18 (59.42)	236.77 (236.77)	254.37 (106.62)	409.85 (409.85)	0.18 (0.18)	0.39 (0.49)	0.12 (0.32)
		Subjunctive	218.88 (64.69)	240.26 (240.26)	265.69 (120.49)	362.40 (362.4)	0.23 (0.23)	0.34 (0.48)	0.13 (0.34)
	Directive	Indicative	219.34 (62.78)	258.53 (258.53)	292.69 (162.35)	435.56 (435.56)	0.16 (0.16)	0.34 (0.48)	0.13 (0.34)
		Subjunctive	216.83 (63.92)	248.67 (248.67)	281.84 (139.45)	430.65 (430.65)	0.17 (0.17)	0.28 (0.45)	0.14 (0.35)
	Emotive–Factive	Indicative	218.26 (66.64)	257.36 (257.36)	282.42 (123.22)	444.26 (444.26)	0.21 (0.21)	0.40 (0.49)	0.13 (0.34)
		Subjunctive	222.60 (71.36)	262.28 (262.28)	294.78 (138.89)	430.72 (430.72)	0.21 (0.21)	0.38 (0.49)	0.15 (0.36)

Table 5. *Cont.*

			FFD	GD	GOPAST	TRT	SKIP	IN	OUT
Critical Region	Desire	Indicative	256.25 (116.47)	272.31 (272.31)	274.14 (152.6)	372.97 (372.97)	0.73 (0.73)	0.62 (0.49)	0.03 (0.17)
		Subjunctive	226.05 (59.21)	230.03 (230.03)	233.1 (65.73)	279.82 (279.82)	0.62 (0.62)	0.16 (0.37)	0.08 (0.27)
	Directive	Indicative	250.81 (97.58)	272.96 (272.96)	278.39 (137.86)	360.99 (360.99)	0.66 (0.66)	0.44 (0.50)	0.06 (0.24)
		Subjunctive	242.43 (81.4)	257.63 (257.63)	259.17 (102.12)	320.7 (320.7)	0.47 (0.47)	0.16 (0.36)	0.03 (0.18)
	Emotive–Factive	Indicative	257.55 (102.8)	269.06 (269.06)	275.14 (129.3)	330.09 (330.09)	0.65 (0.65)	0.47 (0.5)	0.05 (0.21)
		Subjunctive	243.71 (91.26)	255.05 (255.05)	266.03 (126.49)	319.11 (319.11)	0.51 (0.51)	0.17 (0.38)	0.09 (0.29)
Spillover Region	Desire	Indicative	242.42 (86.38)	406.18 (406.18)	511.31 (309.96)	662.7 (662.7)	0.11 (0.11)	0.23 (0.42)	0.41 (0.49)
		Subjunctive	243.30 (73.31)	398.73 (398.73)	426.96 (241.36)	513.76 (513.76)	0.08 (0.08)	0.15 (0.36)	0.21 (0.41)
	Directive	Indicative	249.58 (97.00)	401.14 (401.14)	498.23 (300.01)	669.49 (669.49)	0.11 (0.11)	0.24 (0.43)	0.29 (0.45)
		Subjunctive	230.35 (77.12)	335.40 (335.40)	368.2 (210.71)	501.95 (501.95)	0.13 (0.13)	0.24 (0.43)	0.17 (0.37)
	Emotive–Factive	Indicative	240.90 (75.43)	360.22 (360.22)	420.12 (242.24)	577.22 (577.22)	0.16 (0.16)	0.22 (0.42)	0.32 (0.47)
		Subjunctive	246.49 (86.60)	356.63 (356.63)	383.97 (244.16)	481.63 (481.63)	0.18 (0.18)	0.15 (0.36)	0.17 (0.38)

First fixation duration (FFD); gaze duration (GD); go-past time (GOPAST); total reading time (TRT); skipping probability (SKIP); regression-in probability (IN); regression-out probability (OUT).

Matrix Predicate

The estimates from the mixed-effects models did not reveal any significant effects of interest at the matrix predicate, with the exception of trial number for total reading time (β = 0.00, SE = 0.00, $t = -3.28$, $p < 0.001$), suggesting that this measure decreased as a function of trial number.

Critical Region

At the critical region, there was a significant main effect of grammaticality (β = −0.94, SE = 0.27, z = −3.46, $p < 0.001$; i.e., L1 speakers were more likely to regress back to the critical region for later parts of the sentence) and of trial number (β = −0.01, SE = 0.00, z = −2.07, $p = 0.04$; i.e., regressions back to the critical region were less likely as a function of trial number) for regression-in probability. We additionally found a significant main effect of trial number (β = 0.00, SE = 0.00, $t = -2.01$, $p = 0.04$; i.e., total reading time decreased as a function of proficiency) and a significant interaction between grammaticality and matrix semantic property (desire vs. directive; β = −0.06, SE = 0.03, $t = -2.18$, $p = 0.03$) for total reading time. However, further pairwise analyses of this interaction did not reveal any significant differences between individual conditions.

Spillover Region

At the spillover region, we observed a significant effect of grammaticality (i.e., subjunctive sentences received longer reading times or higher probabilities) for go-past time (β = −0.08, SE = 0.01, $t = -5.72$, $p < 0.001$), total reading time (β = −0.11, SE = 0.02, $t = -7.02$, $p < 0.001$), regression-in probability (β = −0.16, SE = 0.08, z = −1.97, $p = 0.05$), and regression-out probability (β = −0.41, SE = 0.08, z = −5.39, p < 0.001) and a significant effect of trial number (i.e., reading measures decreased as a function of trial number) for go-past time (β = 0.00, SE = 0.00, t = −2.96, p < 0.001), total reading time (β = 0.00, SE = 0.00, t = −4.37, p < 0.001), and regression-in probability (β = −0.01, SE = 0.00, z = −2.57, p < 0.001). Finally, for regression-out probability, we detected a significant effect of matrix semantic property (desire vs. directive; β = 0.27, SE = 0.10, z = 2.29, p = 0.02), such that speakers were more likely to regress back to earlier regions of the sentence with directive than desire predicates.

Summary

In summary, these results revealed that L1 speakers were sensitive to the mood–modality mismatches, but that this sensitivity surfaced predominately at the spillover region in the following measures: go-past time, total reading time, regression-in probability, and regression-out probability. There was, however, some evidence of sensitivity at the critical region via regression-in probability.

5.2.2. L2 Speakers

The raw means and the mixed-effects models for the reading measures are presented in Table 6 below and Table A6 in Appendix A, respectively.

Table 6. Reading time measures for L2 speakers.

			FFD	GD	GOPAST	TRT	SKIP	IN	OUT
Matrix Predicate	Desire	Indicative	231.75 (78.32)	266.97 (266.97)	312.76 (156.75)	525.35 (525.35)	0.25 (0.25)	0.55 (0.50)	0.19 (0.39)
		Subjunctive	230.54 (80.1)	265.46 (265.46)	312.56 (143.00)	507.58 (507.58)	0.25 (0.25)	0.52 (0.50)	0.22 (0.41)
	Directive	Indicative	241.33 (81.12)	278.83 (278.83)	360.82 (201.99)	609.85 (609.85)	0.17 (0.17)	0.49 (0.50)	0.27 (0.44)
		Subjunctive	247.33 (87.37)	307.05 (307.05)	372.96 (204.17)	575.17 (575.17)	0.20 (0.20)	0.45 (0.50)	0.25 (0.43)
	Emotive–Factive	Indicative	243.60 (96.97)	296.25 (296.25)	342.55 (168.14)	582.38 (582.38)	0.25 (0.25)	0.50 (0.50)	0.19 (0.39)
		Subjunctive	233.27 (84.01)	279.19 (279.19)	331.46 (171.68)	577.55 (577.55)	0.24 (0.24)	0.55 (0.50)	0.20 (0.40)
Critical Region	Desire	Indicative	257.34 (97.73)	264.63 (264.63)	274.61 (120.91)	378.85 (378.85)	0.62 (0.62)	0.44 (0.50)	0.06 (0.24)
		Subjunctive	273.65 (107.56)	286.41 (286.41)	294.08 (143.37)	429.48 (429.48)	0.48 (0.48)	0.30 (0.46)	0.08 (0.28)
	Directive	Indicative	255.49 (96.18)	260.75 (260.75)	274.96 (131.28)	401.13 (401.13)	0.62 (0.62)	0.46 (0.50)	0.09 (0.28)
		Subjunctive	274.61 (94.93)	296.33 (296.33)	306.21 (121.77)	451.63 (451.63)	0.39 (0.39)	0.26 (0.44)	0.05 (0.23)
	Emotive–Factive	Indicative	245.53 (82.52)	259.16 (259.16)	274.13 (127.14)	377.85 (377.85)	0.58 (0.58)	0.35 (0.48)	0.13 (0.34)
		Subjunctive	261.92 (95.08)	278.03 (278.03)	294.92 (155.81)	444.68 (444.68)	0.39 (0.39)	0.33 (0.47)	0.12 (0.33)
Spillover Region	Desire	Indicative	268.20 (99.71)	476.27 (476.27)	595.51 (357.39)	907.55 (907.55)	0.07 (0.07)	0.38 (0.49)	0.30 (0.46)
		Subjunctive	268.77 (97.57)	501.93 (501.93)	588.08 (293.33)	857.81 (857.81)	0.06 (0.06)	0.34 (0.47)	0.21 (0.41)

Table 6. *Cont.*

			FFD	GD	GOPAST	TRT	SKIP	IN	OUT
Spillover Region	Directive	Indicative	268.09 (100.11)	460.47 (460.47)	575.12 (401.91)	964.59 (964.59)	0.08 (0.08)	0.41 (0.49)	0.23 (0.42)
		Subjunctive	276.01 (100.99)	456.08 (456.08)	525.59 (330.33)	848.68 (848.68)	0.06 (0.06)	0.36 (0.48)	0.17 (0.38)
	Emotive–Factive	Indicative	266.02 (102.64)	413.42 (413.42)	499.09 (316.46)	764.65 (764.65)	0.12 (0.12)	0.31 (0.46)	0.24 (0.43)
		Subjunctive	265.22 (96.42)	464.18 (464.18)	576.7 (396.22)	848.34 (848.34)	0.07 (0.07)	0.25 (0.43)	0.29 (0.45)

First fixation duration (FFD); gaze duration (GD); go-past time (GOPAST); total reading time (TRT); skipping probability (SKIP); regression-in probability (IN); regression-out probability (OUT).

Matrix Predicate

At the matrix predicate, there was a significant main effect of proficiency (i.e., reading times became quicker as a function of proficiency) for first fixation duration ($\beta = -0.06$, SE = 0.03, $t = -2.04$, $p = 0.05$) and gaze duration ($\beta = -0.08$, SE = 0.04, $t = -2.31$, $p = 0.03$) and of trial number (i.e., reading times and/or probability decreases as a function of trial number) for go-past time ($\beta = 0.00$, SE = 0.00, $t = -2.99$, $p < 0.001$), total reading time ($\beta = 0.00$, SE = 0.00, $t = -3.98$, $p > 0.001$), and regression-in probability ($\beta = -0.01$, SE = 0.00, z = -2.63, $p = 0.01$).

Critical Region

At the critical region, we observed a significant interaction between grammaticality and matrix semantic property (emotive–factive vs. directive; $\beta = 0.33$, SE = 0.16, z = 2.02, $p = 0.04$) for regression-in probability. However, pairwise comparisons did not reveal any significant effects between individual conditions. We further found a significant main effect of frequency for first fixation duration ($\beta = -0.08$, SE = 0.04, $t = -2.16$, $p = 0.04$; i.e., reading times were shorter with higher frequency verbs).

Spillover Region

At the spillover region, we found a significant effect of grammaticality for gaze duration ($\beta = 0.05$, SE = 0.02, $t = 1.96$, $p = 0.05$; i.e., reading times were longer for subjunctive (grammatical) than indicative (ungrammatical) sentences, contra the L1 speakers); of trial number (i.e., reading times and/or probability decreased as a function of trial number) for gaze duration ($\beta = 0.00$, SE = 0.00, $t = -2.71$, $p = 0.01$), total reading time ($\beta = 0.00$, SE = 0.00, $t = -3.56$, $p < 0.001$), and regression-out probability ($\beta = -0.01$, SE = 0.00, z = -2.33, $p = 0.02$), and of proficiency for regression-out probability ($\beta = -0.24$, SE = 0.11, z = -2.11, $p = 0.04$).

Summary

We did not find any evidence to suggest that L2 learners were sensitive to mood–modality mismatches. In fact, L2 speakers exhibited a preference for the indicative over the subjunctive, as evidenced by longer gaze duration for subjunctive than indicative sentences in the spillover region.

6. Discussion

Data from an acceptability judgment task and eye-tracking during reading experiment revealed a number of findings relating to both L1 and L2 speakers' use of the subjunctive in obligatory contexts. In this section, we discuss these findings in relation to the acquisition task predictions outlined in Section 2.3 and to wider debates in the field of L2 acquisition.

Let us first turn our attention to the acceptability judgment ratings (i.e., the "offline" data). Here, we found that the L1 and L2 speakers accurately and reliably differentiated between acceptable (subjunctive) and unacceptable (indicative) sentences, with L2 speakers' ratings becoming more categorical as a function of proficiency. Such findings suggest that both groups had knowledge of the obligatory subjunctive that coincided with the expected norms, with proficiency further strengthening this knowledge among L2 speakers. These findings are therefore consistent with previous research on the L2 acquisition of the French subjunctive (see, e.g., Ayoun 2013; Bartning et al. 2012; Howard 2008; McManus and Mitchell 2015).

In contrast, the eye movement patterns (i.e., the "online" data) painted an entirely different picture. Initially, we anticipated that the L1 speakers would show sensitivity to the mood–modality mismatches at the critical region when reading for comprehension. This did not turn out to be the case, with the exception of regression-in probability. In fact, most of the sensitivity effects were concentrated at the spillover region for go-past time, total reading time, regression-in probability, and regression-out probability.

We must, however, acknowledge at this point that knowledge of the subjunctive is known to stabilise late in (monolingual) L1 acquisition, as is the case in L2 acquisition. For example, early research by Pérez-Leroux (1998, p. 586) on L1 child Spanish found that although monolingual children start to use subjunctive morphosyntax as early as 2;0, "[mood] selection [. . .] is a process that spans over a period of six or seven years". Similarly, Flores et al. (2017) showed that even between the ages of 8 and 9, knowledge of the subjunctive in European Portuguese was not always target-like.

Furthermore, based on findings from a study on delayed grammatical acquisition in L1 development, Meisel et al. (2011) argued that knowledge of grammatical structures (e.g., subject–verb inversion and subject clitics in French interrogatives) that are acquired after the age of 5 may be comparable to that of L2 learners, potentially due to their complexity and relative infrequency in colloquial language. Given the morphosyntactic complexity and relative infrequency of the subjunctive in colloquial language, it is plausible that the late L1 acquisition of the subjunctive may have had a subsequent impact on the strength of L1 speakers' sensitivity, which would thus explain why the sensitivity effects observed in the current study were delayed and not as robust as we initially anticipated.[13]

It is important to advocate caution, here; it is possible that such findings do not necessarily indicate delayed sensitivity among L1 speakers, but instead reflect a by-product (and, in a sense, a limitation of the current study) of the critical region's length. In other words, the critical region was simply too short to detect any significant effects of grammaticality. Regardless of where we found sensitivity effects, the current study showed that the experimental manipulation had an effect on L1 speakers' processing patterns. As an anonymous reviewer suggests, it is equally possible that the native speakers (and the L2 learners) became aware of the experimental manipulation, given that the test items had the same form throughout, which in turn reduced any "surprise" effect upon encountering the indicative where the subjunctive may have been expected. It is for this reason that we included trial number in the analysis to account for such an effect. However, we acknowledge that the fact that participants may have become aware of the experimental manipulation is an evident limitation of the current study.

Such online sensitivity effects, however, were not as present among the L2 speakers, however. Earlier in this article, we referenced a highly pertinent quote from VanPatten and Jegerski (2010, p. 8) stating that the acquisition of the relevant grammatical knowledge "does not automatically entail the processing of something related". The current study provided clear evidence of this being the case, as illustrated by a distinct asymmetry between L2 learners' offline and online knowledge of the subjunctive. Specifically, we found no evidence to suggest that L2 learners could accurately process mood–modality

[13] It is likely that the L1 behaviour could also be attributed to the fact that these speakers were bilinguals. To further explore the validity of this claim, it would be necessary to replicate this study with monolingual L1 speakers of French. However, English is a compulsory subject in French schools and in many universities, which means it is becoming increasingly more challenging to recruit this demographic and explore such questions.

mismatches (i.e., display significantly different reading times between ungrammatical and grammatical sentences) during real-time processing in the L2.

What do these findings mean, then, in terms of our previously formulated research questions and subsequent predictions? First, they show that L2 learners can acquire target-like knowledge of the obligatory subjunctive, regardless of proficiency and/or the lexical–semantic properties of the matrix predicate. However, we found no evidence to suggest that the L1 grammar[14] moderates L2 speakers' offline knowledge of the subjunctive, at least at the proficiency levels tested, contra our initial predictions.[15] In fact, our learners may have been too advanced for us to have detected the feature-based distinctions among the selecting predicates. A likely factor in this successful acquisition is the explicit instruction of the subjunctive in French language classrooms.

What our study did reveal, however, is that this (offline) knowledge did not translate into target-like processing patterns. In other words, the L2 speakers in our study were unable to apply their offline knowledge during real-time processing to detect morphosyntactic violations, at least at the proficiency levels tested.[16] Furthermore, and contra our initial predictions, there was no evidence to suggest that the lexical–semantic properties of the matrix predicate, and, by extension, the L1, moderated this (lack of) sensitivity. One possible explanation for these findings is that learners may have already acquired the target-like featural specifications of the obligatory subjunctive, but that delayed lexical access and/or limited computational resources prevented learners from applying this knowledge during real-time processing. Such an argument is espoused most notably by limited capacity models (e.g., McDonald 2006; Hopp 2010; Sorace 2011; Dekydtspotter and Renaud 2014).

An alternative explanation for such findings references the exact nature of the knowledge that L2 learners demonstrated in the acceptability judgment task. In other words, does this knowledge necessarily reflect "an intuitive, automatic type of linguistic knowledge ('acquired' or 'implicit' knowledge)", or rather "a deliberate, consciously controlled type of language termed 'learned' or 'explicit' knowledge)" (Whong et al. 2014, p. 553)? We argue that the knowledge observed in our study is most likely to reflect the latter (i.e., "learned" or "explicit" knowledge), since in an acceptability judgment task, learners can (and are, in a sense, encouraged to) access metalinguistic knowledge. This in turn increases the likelihood of instruction effects, especially with properties, such as the subjunctive, which are highly taught in the classroom setting. Therefore, it is likely that these instruction effects obscure L1 transfer effects, which would in turn explain why the current findings do not fully support our initial predictions. For further discussion on implicit and explicit learning in relation to the role of instruction in the L2 acquisition of the Spanish subjunctive, see van Osch et al. (2018).

Conversely, in an online task, such as eye-tracking during reading, lexical access and syntactic computation are expected to occur quickly and seamlessly, meaning that L2 learners do not have time to access their metalinguistic knowledge of language and thus must rely on an intuitive, automatic type of linguistic knowledge. The discrepancy that we found between the offline and online data therefore speaks to the type of knowledge

---

14 An anonymous reviewer suggests that it could be the case that British English L2 learners of French do not establish a relation between the subjunctive in English and French. However, the Feature Reassembly Hypothesis conceptualises linguistic structures as feature bundles, rather than isolated structures. As such, it is not a question of whether learners consciously establish a relation between the subjunctive in English and French, but rather whether the structure shares the same features across the two languages.

15 Although we based our predicate-by-predicate hypotheses on an expected L1 influence, it is possible that the subjunctive with directive verbs is generally learned earlier than desire verbs, irrespective of the L1, given that Bartning and Schlyter (2004) report that *il faut que* is acquired relatively early by Swedish L1 learners of French. When investigating a possible L1 influence, future studies should consider comparing L2 learners from at least two L1 to determine whether the L2 behaviour is the result of L1 influence or not.

16 An anonymous reviewer pointed out that the L2 judgments for the indicative were higher than the L1 judgments for the indicative, and asked whether the difference was significant. In our study, and in a bid to avoid the comparative fallacy, we were not interested in whether there was a significant difference between the L1 and L2 speakers, but rather whether each group was able to significantly differentiate between grammatical and ungrammatical sentences.

that L2 learners have: knowledge about the subjunctive may be more learned knowledge rather than internalised representation.

We must not, however, ignore the wider theoretical implications of such findings. As we discussed in the introduction, the main aim of the current paper was to explore whether a processing-based extension of the Feature Reassembly Hypothesis was theoretically and empirically viable. The evidence presented in this article, albeit limited, suggests that the predictions of the FRH do not hold true for processing patterns, at least in the context of the structure and proficiency levels tested in this study. This is understandable given the additional cognitive demands that real-time processing places on the parser. It is important that future research directly examines the nature of L2 computational resources, by exploring how measures of working memory, for example, correlate with L2 processing patterns. Furthermore, it would be interesting to replicate the current study, by undertaking further testing with different populations, such as highly advanced to near-native L2 speakers of French with extensive exposure to the target language, in order to further explore the role of proficiency, or L2 speakers with relatively limited to no exposure to classroom instruction, in order to explore the effect of explicit instruction. Similarly, given how highly taught the subjunctive is in both the L1 and L2 classroom, it would be worthwhile to investigate subjunctive knowledge among heritage speakers who are, in large part, unlikely to have received extensive explicit instruction about the French subjunctive in order to address how learning context modulates linguistic development.

While the current findings may not necessarily support a processing-based extension of the Feature Reassembly Hypothesis, further discussion of how our findings relate to other theories, such as the Shallow Structure Hypothesis, is necessary. For example, under a Shallow Structure Hypothesis interpretation of the current findings, it could be argued that L2 learners' lack of online sensitivity to mood–modality mismatches demonstrates their difficulty applying syntactic information or lexical access difficulties relating to the relative infrequency of the subjunctive. However, if this was indeed the case, we should have found evidence that L2 learners prioritised semantic information by regressing back to the matrix predicate, where the semantic information is expressed, rather than slowing down at the critical region, where the morphosyntactic information is expressed. This, however, was not the case, as there were no significant main effects of grammaticality between indicative and subjunctive sentences at the matrix predicate, or at the critical region for that matter.

In conclusion, the current study found a distinct asymmetry between L2 learners' knowledge of the subjunctive and their subsequent processing patterns. Specifically, while L2 (offline) knowledge of the subjunctive was consistently target-like, this was distinctly not the case for processing patterns. These findings are compatible with findings from a wide variety of L2 properties (see, e.g., Papadopoulou and Clahsen 2003 for relative clause attachment in Greek; Marinis et al. 2005 for long-distance wh-dependencies in English; Shimanskaya and Slabakova 2017 for pronominal clitics in French; Stepanov et al. 2019 for long-distance syntactic dependencies in English). Contra our initial predictions, the L1 did not appear to modulate sensitivity to mood–modality mismatches during real-time processing. We argued that such findings have two possible interpretations: (1) that knowledge of the subjunctive is "learned" knowledge, rather than "internalised" representations and (2) that reduced lexical access and/or limited computational resources (e.g., working memory) prevented learners from fully utilising their grammatical representations during real-time processing.

Author Contributions: Investigation, A.D.; Supervision, R.S.; Writing—original draft, A.D.; Writing—review & editing, R.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the UK Economic and Social Research Council (ESRC), grant number ES/J500161/1 to James Corbet.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by the Ethics Committee of the University of Southampton (ERGO #46014).

Informed Consent Statement: All subjects gave their informed consent for inclusion before they participated in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. Participants' age of acquisition and self-assessed proficiency scores.

			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)

Table A2. 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	6	9
	Female	24	36
Education	Undergraduate	23	40
	Postgraduate (Masters)	7	5
	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)

Table A2. *Cont.*

	L1 Speakers (*n* = 30)	L2 Speakers (*n* = 45)
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)

Table A3. Estimates from the mixed-effect model for L1 judgment data.

	β	*SE*	*t*	*p*
(Intercept)	4.87	0.18	27.04	<0.001
Mood: Subjunctive vs. Indicative	1.94	0.03	58.48	<0.001
Semantic: Desire vs. Directive	0.01	0.12	0.07	0.95
Semantic: Emotive–Factive vs. Directive	−0.02	0.12	−0.13	0.90
Sentence Length (Centred)	0.15	0.09	1.68	0.10
Trial Number	0.00	0.00	−1.43	0.15
Mood × Semantic (Desire vs. Directive)	0.03	0.05	0.75	0.45
Mood × Semantic (Emotive–Factive vs. Directive)	0.01	0.05	0.14	0.89

Table A4. Estimates from the mixed-effect model for L2 judgment data.

	b	*SE*	*t*	*p*
(Intercept)	5.45	0.22	24.67	<0.001
Mood: Subjunctive vs. Indicative	1.23	0.06	19.44	<0.001
Semantic: Desire vs. Directive	−0.16	0.10	−1.69	0.09
Semantic: Emotive–Factive vs. Directive	−0.03	0.09	−0.35	0.73
Proficiency (Centred)	−0.29	0.24	−1.24	0.22
Sentence Length (Centred)	−0.06	0.04	−1.37	0.18
Trial Number	−0.01	0.00	−2.91	<0.001
Mood × Semantic (Desire vs. Directive)	0.15	0.09	1.68	0.09
Mood × Semantic (Emotive–Factive vs. Directive)	−0.04	0.09	−0.43	0.67
Mood × Proficiency	0.53	0.07	7.29	<0.001
Semantic (Desire vs. Directive) × Prof.	−0.01	0.10	−0.10	0.92
Semantic (Emotive–Factive vs. Directive) × Prof.	−0.13	0.10	−1.27	0.20
Mood × Semantic (Desire vs. Directive) × Prof.	−0.04	0.10	−0.37	0.71
Mood × Semantic (Emotive–Factive vs. Directive) × Prof.	−0.01	0.10	−0.13	0.90

Table A5. Estimates from mixed-effects models for L1 eye movement data.

		Matrix Predicate				Critical Region				Spillover Region			
		b	***SE***	***t*/*z***	***p***	***b***	***SE***	***t*/*z***	***p***	***b***	***SE***	***t*/*z***	***p***
FFD	(Intercept)	5.35	0.02	249.89	<0.001	5.41	0.03	155.26	<0.001	5.44	0.03	206.67	<0.001
	Mood: Subjunctive vs. Indicative	0.00	0.01	0.36	0.72	−0.08	0.04	−2.03	0.05	−0.01	0.01	−0.80	0.43
	Semantic: Desire vs. Directive	−0.01	0.01	−0.39	0.70	−0.02	0.02	−0.85	0.40	0.01	0.02	0.35	0.73
	Semantic: Emotive–Factive vs. Directive	0.01	0.01	0.39	0.70	0.02	0.02	1.15	0.26	0.01	0.02	0.34	0.74
	Region Frequency (Centred)	-	-	-	-	−0.06	0.04	−1.42	0.17	-	-	-	-
	Trial Number	0.00	0.00	0.14	0.89	0.00	0.00	0.92	0.36	0.00	0.00	0.10	0.92
	Mood × Semantic (Desire vs. Directive)	0.01	0.01	0.52	0.61	−0.03	0.02	−1.47	0.14	0.01	0.01	1.13	0.26
	Mood × Semantic (Emotive–Factive vs. Directive)	0.00	0.01	0.44	0.66	0.01	0.02	0.31	0.76	0.01	0.01	1.03	0.30
GD	(Intercept)	5.46	0.03	178.62	<0.001	5.44	0.04	141.24	<0.001	5.80	0.05	111.92	<0.001
	Mood: Subjunctive vs. Indicative	0.00	0.01	−0.21	0.84	−0.09	0.05	−1.87	0.07	−0.03	0.02	−1.38	0.18
	Semantic: Desire vs. Directive	−0.03	0.02	−1.42	0.17	−0.03	0.02	−1.25	0.22	0.07	0.05	1.29	0.21
	Semantic: Emotive–Factive vs. Directive	0.03	0.02	1.17	0.25	0.02	0.02	0.81	0.43	−0.06	0.05	−1.16	0.25
	Region Frequency (Centred)	-	-	-	-	−0.06	0.04	−1.47	0.15	-	-	-	-
	Trial Number	0.00	0.00	−0.37	0.72	0.00	0.00	1.20	0.23	0.00	0.00	−1.00	0.32
	Mood × Semantic (Desire vs. Directive)	0.00	0.02	0.20	0.85	−0.03	0.02	−1.40	0.17	0.02	0.03	0.74	0.46
	Mood × Semantic (Emotive–Factive vs. Directive)	0.01	0.02	0.54	0.59	0.01	0.02	0.51	0.62	0.02	0.03	0.88	0.39
GOPAST	(Intercept)	5.57	0.04	151.48	<0.001	5.47	0.04	136.80	<0.001	5.98	0.06	99.57	<0.001
	Mood: Subjunctive vs. Indicative	0.01	0.01	0.60	0.55	−0.08	0.05	−1.61	0.12	−0.08	0.01	−5.72	<0.001
	Semantic: Desire vs. Directive	−0.05	0.03	−1.92	0.06	−0.03	0.02	−1.32	0.20	0.08	0.07	1.20	0.24
	Semantic: Emotive–Factive vs. Directive	0.04	0.03	1.34	0.19	0.02	0.02	1.04	0.31	−0.08	0.07	−1.21	0.23
	Region Frequency (Centred)	-	-	-	-	−0.06	0.05	−1.29	0.21	-	-	-	-
	Trial Number	0.00	0.00	−1.41	0.16	0.00	0.00	0.56	0.57	0.00	0.00	−2.96	<0.001
	Mood × Semantic (Desire vs. Directive)	0.01	0.02	0.42	0.68	−0.03	0.02	−1.36	0.17	0.01	0.02	0.53	0.60
	Mood × Semantic (Emotive–Factive vs. Directive)	0.01	0.02	0.85	0.40	0.02	0.02	0.89	0.38	0.04	0.02	1.88	0.06
TRT	(Intercept)	5.96	0.05	113.47	<0.001	5.69	0.05	113.08	<0.001	6.26	0.07	92.00	<0.001
	Mood: Subjunctive vs. Indicative	−0.02	0.02	−1.38	0.17	−0.10	0.05	−1.80	0.07	−0.11	0.02	−7.02	<0.001
	Semantic: Desire vs. Directive	−0.06	0.04	−1.53	0.14	−0.03	0.03	−1.05	0.29	0.05	0.06	0.82	0.42
	Semantic: Emotive–Factive vs. Directive	0.03	0.04	0.79	0.44	−0.01	0.03	−0.29	0.77	−0.08	0.06	−1.29	0.21
	Region Frequency (Centred)	-	-	-	-	−0.05	0.05	−0.91	0.36	-	-	-	-
	Trial Number	0.00	0.00	−3.28	<0.001	0.00	0.00	−2.01	0.04	0.00	0.00	−4.37	<0.001
	Mood × Semantic (Desire vs. Directive)	−0.03	0.02	−1.20	0.23	−0.06	0.03	−2.18	0.03	0.00	0.02	−0.03	0.97
	Mood × Semantic (Emotive–Factive vs. Directive)	0.02	0.02	0.94	0.35	0.04	0.03	1.53	0.13	0.03	0.02	1.15	0.25

Table A5. *Cont.*

		Matrix Predicate				Critical Region				Spillover Region			
		b	***SE***	***t/z***	***p***	***b***	***SE***	***t/z***	***p***	***b***	***SE***	***t/z***	***p***
IN	(Intercept)	−0.45	0.16	−2.77	0.01	−0.48	0.22	−2.20	0.03	−1.14	0.19	−5.94	<0.001
	Mood: Subjunctive vs. Indicative	−0.10	0.07	−1.47	0.14	−0.94	0.27	−3.46	<0.001	−0.16	0.08	−1.97	0.05
	Semantic: Desire vs. Directive	0.04	0.10	0.45	0.65	0.23	0.15	1.54	0.12	−0.08	0.12	−0.71	0.48
	Semantic: Emotive–Factive vs. Directive	0.15	0.10	1.55	0.12	−0.06	0.14	−0.40	0.69	−0.15	0.12	−1.27	0.21
	Region Frequency (Centred)	-	-	-	-	−0.06	0.26	−0.24	0.81	-	-	-	-
	Trial Number	−0.01	0.00	−1.52	0.13	−0.01	0.00	−2.07	0.04	−0.01	0.00	−2.57	0.01
	Mood × Semantic (Desire vs. Directive)	−0.01	0.10	−0.07	0.94	−0.20	0.15	−1.37	0.17	−0.09	0.12	−0.80	0.42
	Mood × Semantic (Emotive–Factive vs. Directive)	0.05	0.10	0.55	0.58	0.10	0.14	0.72	0.47	−0.06	0.12	−0.52	0.60
OUT	(Intercept)	−1.71	0.22	−7.78	<0.001	−2.84	0.36	−7.97	<0.001	−0.97	0.15	−6.60	<0.001
	Mood: Subjunctive vs. Indicative	0.07	0.10	0.78	0.43	0.04	0.51	0.07	0.95	−0.41	0.08	−5.39	<0.001
	Semantic: Desire vs. Directive	−0.10	0.14	−0.73	0.46	−0.11	0.28	−0.39	0.70	0.24	0.10	2.29	0.02
	Semantic: Emotive–Factive vs. Directive	0.05	0.13	0.39	0.70	0.27	0.26	1.07	0.29	−0.07	0.11	−0.66	0.51
	Region Frequency (Centred)	-	-	-	-	−0.16	0.50	−0.32	0.75	-	-	-	-
	Trial Number	−0.01	0.00	−1.90	0.06	0.00	0.01	−0.16	0.87	0.00	0.00	−1.00	0.32
	Mood × Semantic (Desire vs. Directive)	−0.01	0.14	−0.04	0.97	0.31	0.28	1.12	0.26	−0.07	0.10	−0.71	0.48
	Mood × Semantic (Emotive–Factive vs. Directive)	0.04	0.13	0.28	0.78	0.18	0.25	0.71	0.48	0.02	0.11	0.14	0.89

Table A6. Estimates from mixed-effects models for L2 eye movement data.

		Matrix Predicate				Critical Region				Spillover Region			
		b	***SE***	***t/z***	***p***	***b***	***SE***	***t/z***	***p***	***b***	***SE***	***t/z***	***p***
FFD	(Intercept)	5.38	0.03	183.38	<0.001	5.49	0.04	153.32	<0.001	5.56	0.03	177.46	<0.001
	Mood: Subjunctive vs. Indicative	−0.02	0.01	−1.55	0.12	−0.05	0.04	−1.25	0.22	0.00	0.01	−0.02	0.99
	Semantic: Desire vs. Directive	−0.01	0.02	−0.46	0.64	0.02	0.03	0.78	0.44	0.00	0.02	−0.06	0.95
	Semantic: Emotive–Factive vs. Directive	0.00	0.02	0.11	0.91	−0.02	0.03	−0.58	0.56	0.00	0.02	0.03	0.98
	Proficiency (Centred)	−0.06	0.03	−2.04	0.05	−0.01	0.03	−0.35	0.73	0.00	0.03	0.00	1.00
	Region Frequency (Centred)	-	-	-	-	−0.08	0.04	−2.16	0.04	-	-	-	-
	Trial Number	0.00	0.00	−0.35	0.72	0.00	0.00	0.64	0.52	0.00	0.00	−1.58	0.12
	Mood × Semantic (Desire vs. Directive)	−0.01	0.02	−0.77	0.44	0.02	0.02	0.93	0.35	0.01	0.02	0.59	0.55
	Mood × Semantic (Emotive–Factive vs. Directive)	0.01	0.02	0.45	0.65	−0.01	0.02	−0.53	0.60	−0.01	0.02	−0.44	0.66
	Mood × Proficiency	−0.03	0.02	−1.71	0.09	−0.01	0.02	−0.55	0.58	−0.01	0.02	−0.48	0.63
	Semantic (Desire vs. Directive) × Prof.	0.02	0.02	1.05	0.29	0.01	0.03	0.26	0.80	0.00	0.02	−0.19	0.85
	Semantic (Emotive–Factive vs. Directive) × Prof.	0.01	0.02	0.28	0.78	0.00	0.03	0.10	0.92	0.02	0.02	0.78	0.43
	Mood × Semantic (Desire vs. Directive) × Prof.	−0.02	0.02	−0.97	0.33	0.04	0.03	1.40	0.16	0.02	0.02	0.96	0.34
	Mood × Semantic (Emotive–Factive vs. Directive) × Prof.	0.03	0.02	1.38	0.17	−0.01	0.03	−0.49	0.62	0.00	0.02	−0.05	0.96
GD	(Intercept)	5.51	0.04	146.24	<0.001	5.51	0.04	134.83	<0.001	5.99	0.06	100.50	<0.001
	Mood: Subjunctive vs. Indicative	−0.01	0.02	−0.33	0.74	−0.02	0.04	−0.54	0.59	0.05	0.02	1.96	0.05
	Semantic: Desire vs. Directive	−0.03	0.03	−0.90	0.37	0.02	0.03	0.54	0.59	0.07	0.06	1.12	0.27
	Semantic: Emotive–Factive vs. Directive	0.04	0.03	1.34	0.18	−0.02	0.03	−0.69	0.49	−0.10	0.06	−1.55	0.13
	Proficiency (Centred)	−0.08	0.04	−2.31	0.03	−0.03	0.04	−0.74	0.46	−0.05	0.05	−1.06	0.29
	Region Frequency (Centred)	-	-	-	-	−0.06	0.04	−1.55	0.13	-	-	-	-
	Trial Number	0.00	0.00	−1.00	0.32	0.00	0.00	0.26	0.80	0.00	0.00	−2.71	0.01
	Mood × Semantic (Desire vs. Directive)	−0.03	0.03	−1.20	0.23	0.02	0.03	0.76	0.45	0.01	0.03	0.44	0.66
	Mood × Semantic (Emotive–Factive vs. Directive)	0.01	0.03	0.36	0.72	−0.03	0.03	−1.15	0.25	0.02	0.03	0.69	0.49
	Mood × Proficiency	−0.01	0.02	−0.56	0.58	−0.01	0.02	−0.54	0.59	0.03	0.03	0.97	0.33
	Semantic (Desire vs. Directive) × Prof.	0.03	0.03	0.89	0.37	0.01	0.03	0.47	0.64	0.01	0.04	0.16	0.87
	Semantic (Emotive–Factive vs. Directive) × Prof.	0.03	0.03	1.11	0.27	−0.01	0.03	−0.32	0.75	−0.03	0.04	−0.84	0.40
	Mood × Semantic (Desire vs. Directive) × Prof.	−0.04	0.03	−1.35	0.18	0.04	0.03	1.46	0.15	0.00	0.04	0.12	0.90
	Mood × Semantic (Emotive–Factive vs. Directive) × Prof.	0.05	0.03	1.69	0.09	−0.03	0.03	−0.87	0.38	0.03	0.04	0.86	0.39

Table A6. *Cont.*

		Matrix Predicate				Critical Region				Spillover Region			
		b	***SE***	***t/z***	***p***	***b***	***SE***	***t/z***	***p***	***b***	***SE***	***t/z***	***p***
GOPAST	(Intercept)	5.74	0.05	111.63	<0.001	5.55	0.04	125.26	<0.001	6.12	0.07	85.47	<0.001
	Mood: Subjunctive vs. Indicative	0.01	0.02	0.52	0.60	−0.06	0.04	−1.46	0.15	0.03	0.02	1.44	0.15
	Semantic: Desire vs. Directive	−0.07	0.04	−1.73	0.09	0.01	0.03	0.26	0.80	0.08	0.08	1.06	0.30
	Semantic: Emotive−Factive vs. Directive	−0.01	0.04	−0.25	0.80	−0.02	0.03	−0.65	0.52	−0.09	0.08	−1.17	0.25
	Proficiency (Centred)	−0.04	0.05	−0.72	0.48	−0.02	0.04	−0.40	0.69	−0.09	0.05	−1.77	0.08
	Region Frequency (Centred)	-	-	-	-	−0.08	0.04	−1.95	0.06	-	-	-	-
	Trial Number	0.00	0.00	−2.99	<0.001	0.00	0.00	0.02	0.98	0.00	0.00	−1.68	0.09
	Mood × Semantic (Desire vs. Directive)	0.00	0.03	0.02	0.98	0.02	0.03	0.76	0.45	0.00	0.03	−0.11	0.91
	Mood × Semantic (Emotive–Factive vs. Directive)	0.02	0.03	0.88	0.38	−0.03	0.03	−0.96	0.34	0.03	0.03	0.84	0.40
	Mood × Proficiency	0.01	0.02	0.51	0.61	−0.04	0.02	−1.93	0.05	0.02	0.02	0.74	0.46
	Semantic (Desire vs. Directive) × Prof.	0.00	0.03	−0.01	1.00	0.01	0.03	0.45	0.66	0.00	0.03	−0.01	0.99
	Semantic (Emotive–Factive vs. Directive) × Prof.	−0.01	0.03	−0.30	0.77	−0.02	0.03	−0.63	0.53	−0.04	0.04	−1.00	0.32
	Mood × Semantic (Desire vs. Directive) × Prof.	−0.01	0.03	−0.32	0.75	0.05	0.03	1.54	0.13	0.00	0.03	0.00	1.00
	Mood × Semantic (Emotive–Factive vs. Directive) × Prof.	0.05	0.03	1.48	0.14	-0.03	0.03	-0.87	0.38	0.00	0.04	-0.03	0.98
TRT	(Intercept)	6.17	0.09	66.84	<0.001	5.74	0.07	79.59	<0.001	6.53	0.10	67.00	<0.001
	Mood: Subjunctive vs. Indicative	0.02	0.02	0.79	0.43	−0.08	0.08	−0.98	0.33	0.00	0.02	−0.15	0.88
	Semantic: Desire vs. Directive	−0.09	0.05	−1.66	0.10	−0.02	0.05	−0.43	0.67	0.02	0.09	0.18	0.86
	Semantic: Emotive–Factive vs. Directive	0.01	0.05	0.27	0.79	−0.01	0.05	−0.19	0.85	−0.14	0.09	−1.51	0.14
	Proficiency (Centred)	−0.07	0.10	−0.77	0.44	−0.05	0.07	−0.64	0.53	-	-	-	-
	Region Frequency (Centred)	-	-	-	-	−0.12	0.08	−1.58	0.12	−0.11	0.08	−1.31	0.20
	Trial Number	0.00	0.00	−3.98	<0.001	0.00	0.00	1.63	0.10	0.00	0.00	−3.56	<0.001
	Mood × Semantic (Desire vs. Directive)	0.02	0.03	0.55	0.58	−0.04	0.04	−1.09	0.27	0.00	0.03	0.15	0.88
	Mood × Semantic (Emotive–Factive vs. Directive)	0.03	0.03	0.76	0.45	0.00	0.04	−0.09	0.93	0.05	0.03	1.51	0.13
	Mood × Proficiency	0.04	0.03	1.53	0.13	−0.06	0.03	−1.86	0.06	−0.02	0.03	−0.65	0.51
	Semantic (Desire vs. Directive) × Prof.	−0.01	0.04	−0.39	0.70	0.00	0.04	−0.01	0.99	−0.04	0.04	−0.98	0.33
	Semantic (Emotive–Factive vs. Directive) × Prof.	−0.01	0.04	−0.15	0.88	−0.02	0.04	−0.43	0.67	−0.06	0.04	−1.54	0.12
	Mood × Semantic (Desire vs. Directive) × Prof.	0.02	0.04	0.53	0.60	−0.05	0.04	−1.05	0.29	0.03	0.04	0.69	0.49
	Mood × Semantic (Emotive–Factive vs. Directive) × Prof.	−0.01	0.04	−0.19	0.85	0.02	0.04	0.57	0.57	0.02	0.04	0.47	0.64

Table A6. *Cont.*

		Matrix Predicate				Critical Region				Spillover Region			
		b	*SE*	*t*/*z*	*p*	*b*	*SE*	*t*/*z*	*p*	*b*	*SE*	*t*/*z*	*p*
IN	(Intercept)	0.19	0.25	0.74	0.46	−0.77	0.16	−4.68	<0.001	−2.33	0.23	−10.28	<0.001
	Mood: Subjunctive vs. Indicative	0.03	0.10	0.27	0.79	−0.18	0.21	−0.84	0.40	−0.20	0.16	−1.26	0.21
	Semantic: Desire vs. Directive	0.00	0.14	−0.03	0.97	−0.11	0.17	−0.65	0.52	−0.21	0.23	−0.93	0.35
	Semantic: Emotive–Factive vs. Directive	0.10	0.14	0.69	0.49	0.02	0.16	0.13	0.89	0.35	0.21	1.70	0.09
	Proficiency (Centred)	−0.12	0.26	−0.44	0.66	−0.03	0.13	−0.21	0.84	0.31	0.19	1.68	0.09
	Region Frequency (Centred)	-	-	-	-	0.15	0.19	0.80	0.42	-	-	-	-
	Trial Number	−0.01	0.00	−2.63	0.01	0.00	0.00	1.26	0.21	0.00	0.00	0.13	0.89
	Mood × Semantic (Desire vs. Directive)	−0.05	0.14	−0.39	0.70	−0.19	0.17	−1.12	0.26	0.04	0.23	0.16	0.87
	Mood × Semantic (Emotive–Factive vs. Directive)	−0.04	0.14	−0.25	0.80	0.33	0.16	2.02	0.04	0.05	0.21	0.25	0.80
	Mood × Proficiency	0.07	0.11	0.59	0.56	−0.08	0.13	−0.62	0.54	0.01	0.19	0.07	0.94
	Semantic (Desire vs. Directive) × Prof.	−0.15	0.16	−0.95	0.34	−0.22	0.19	−1.13	0.26	−0.09	0.27	−0.34	0.74
	Semantic (Emotive–Factive vs. Directive) × Prof.	0.03	0.16	0.20	0.84	0.15	0.18	0.79	0.43	0.20	0.25	0.80	0.42
	Mood × Semantic (Desire vs. Directive) × Prof.	−0.02	0.16	−0.10	0.92	−0.21	0.19	−1.07	0.28	−0.12	0.27	−0.43	0.66
	Mood × Semantic (Emotive–Factive vs. Directive) × Prof.	−0.22	0.16	−1.34	0.18	0.16	0.18	0.86	0.39	0.30	0.25	1.20	0.23
OUT	(Intercept)	−1.03	0.15	−6.72	<0.001	−2.41	0.28	−8.74	<0.001	−0.64	0.14	−4.68	<0.001
	Mood: Subjunctive vs. Indicative	0.04	0.11	0.39	0.70	−0.13	0.36	−0.36	0.72	−0.16	0.10	−1.58	0.11
	Semantic: Desire vs. Directive	−0.24	0.16	−1.51	0.13	0.09	0.29	0.32	0.75	0.06	0.14	0.42	0.67
	Semantic: Emotive–Factive vs. Directive	−0.22	0.16	−1.38	0.17	−0.02	0.29	−0.07	0.94	−0.29	0.15	−1.90	0.06
	Proficiency (Centred)	0.24	0.13	1.90	0.06	−0.19	0.23	−0.80	0.42	−0.24	0.11	−2.11	0.04
	Region Frequency (Centred)	-	-	-	-	−0.02	0.31	−0.06	0.95	-	-	-	-
	Trial Number	0.00	0.00	−0.83	0.41	0.00	0.01	−0.66	0.51	−0.01	0.00	−2.33	0.02
	Mood × Semantic (Desire vs. Directive)	0.27	0.16	1.70	0.09	−0.03	0.29	−0.09	0.93	0.11	0.14	0.79	0.43
	Mood × Semantic (Emotive–Factive vs. Directive)	−0.17	0.16	−1.05	0.29	−0.01	0.29	−0.04	0.97	−0.10	0.15	−0.69	0.49
	Mood × Proficiency	0.04	0.13	0.31	0.76	−0.08	0.23	−0.36	0.72	−0.06	0.11	−0.51	0.61
	Semantic (Desire vs. Directive) × Prof.	−0.22	0.18	−1.21	0.23	0.38	0.33	1.13	0.26	−0.06	0.16	−0.36	0.72
	Semantic (Emotive–Factive vs. Directive) × Prof.	−0.12	0.18	−0.63	0.53	−0.58	0.32	−1.84	0.07	0.01	0.17	0.05	0.96
	Mood × Semantic (Desire vs. Directive) × Prof.	0.29	0.18	1.58	0.11	−0.33	0.34	−0.99	0.32	0.11	0.16	0.68	0.49
	Mood × Semantic (Emotive–Factive vs. Directive) × Prof.	−0.25	0.18	−1.34	0.18	0.01	0.32	0.03	0.97	−0.09	0.17	−0.55	0.58

References

Ayoun, Dalila. 2013. *The Second Language Acquisition of French Tense, Aspect, Mood and Modality*. AILA Applied Linguistics Series; Amsterdam: John Benjamins Publishing Company, vol. 10. [CrossRef]

Barr, Dale J., Roger Levy, Christoph Scheepers, and Harry J. Tily. 2013. Random Effects Structure for Confirmatory Hypothesis Testing: Keep It Maximal. *Journal of Memory and Language* 68: 255–78. [CrossRef] [PubMed]

Bartning, Inge, and Suzanne Schlyter. 2004. Itinéraires Acquisitionnels et Stades de Développement En Français L2. *Journal of French Language Studies* 14: 281–99. [CrossRef]

Bartning, Inge, Fanny Forsberg Lundell, and Victorine Hancock. 2012. On the Role of Linguistic Contextual Factors for Morphosyntactic Stabilization in High-Level L2 French. *Studies in Second Language Acquisition* 34: 243–67. [CrossRef]

Bartning, Inge. 2008. The Advanced Learner Variety: 10 Years Later. In *Contemporary Studies in Descriptive Linguistics, Volume 12: Advanced Learner Variety: The Case of French*. Edited by Emmanuelle Labeau and Florence Myles. Oxford: Peter Lang AG, pp. 11–40.

Bates, Douglas, Martin Mächler, Ben Bolker, and Steve Walker. 2015. Fitting Linear Mixed-Effects Models Using Lme4. *Journal of Statistical Software* 67. [CrossRef]

Baum, Shari, and Debra A. Titone. 2014. Moving toward a Neuroplasticity View of Bilingualism, Executive Control, and Aging. *Applied Psycholinguistics* 35: 857–94. [CrossRef]

Baunaz, Léna. 2017. Embedding Verbs and Subjunctive Mood: The Emotive Factor. In *Romance Languages and Linguistic Theory 11. Selected Papers from the 44th Linguistic Symposium on Romance Languages*. Edited by Silvia Perpiñán, David Heap and Itziri Moreno-Villamar. Amsterdam: John Benjamins Publishing Company, pp. 9–31. [CrossRef]

Bianchi, Valentina. 2001. *On Person Agreement*, Mss., Scuole Normale Superior. Pisa.

Bley-Vroman, Robert. 1983. The Comparative Fallacy in Interlanguage Studies: The Case of Systematicity. *Language Learning* 33: 1–17. [CrossRef]

Boxell, Oliver, and Claudia Felser. 2017. Sensitivity to Parasitic Gaps inside Subject Islands in Native and Non-Native Sentence Processing. *Bilingualism: Language and Cognition* 20: 494–511. [CrossRef]

Brysbaert, Marc, Drieghe Denis, and Vitu Françoise. 2005. Word skipping: implications for theories of eye movement control in reading. In *Cognitive Processes in Eye Guidance*. Edited by Geoffrey Underwood. Oxford: Oxford University Press, pp. 53–78.

Brysbaert, Marc. 2013. Lextale_FR A Fast, Free, and Efficient Test to Measure Language Proficiency in French. *Psychologica Belgica* 53: 23. [CrossRef]

Cho, Jacee, and Roumyana Slabakova. 2014. Interpreting Definiteness in a Second Language without Articles: The Case of L2 Russian. *Second Language Research* 30: 159–90. [CrossRef]

Dekydtspotter, Laurent, and Claire Renaud. 2014. On Second Language Processing and Grammatical Development. *Linguistic Approaches to Bilingualism* 4: 131–65. [CrossRef]

Domínguez, Laura, María J. Arche, and Florence Myles. 2017. Spanish Imperfect Revisited: Exploring L1 Influence in the Reassembly of Imperfective Features onto New L2 Forms. *Second Language Research* 33: 431–57. [CrossRef]

Dudley, Amber. 2020. An Investigation into the Acquisition and Processing of the Subjunctive by English-Speaking Second Language Learners of French. Ph.D. thesis, University of Southampton, Southampton, UK.

Felser, Claudia, and Ian Cunnings. 2012. Processing Reflexives in a Second Language: The Timing of Structural and Discourse-Level Constraints. *Applied Psycholinguistics* 33: 571–603. [CrossRef]

Felser, Claudia, Ian Cunnings, Claire Batterham, and Harald Clahsen. 2012. The Timing of Island Effects in Nonnative Sentence Processing. *Studies in Second Language Acquisition* 34: 67–98. [CrossRef]

Flores, Cristina, Ana Lúcia Santos, Alice Jesus, and Rui Marques. 2017. Age and Input Effects in the Acquisition of Mood in Heritage Portuguese. *Journal of Child Language* 44: 795–828. [CrossRef]

Fodor, Janet Dean. 1998a. Leaning to Parse? *Journal of Psycholinguistic Research* 27: 285–319. [CrossRef]

Fodor, Janet Dean. 1998b. Parsing to Learn. *Journal of Psycholinguistic Research* 27: 339–74. [CrossRef]

Franceschina, Florencia. 2005. *Fossilized Second Language Grammars*. Language Acquisition and Language Disorders. Amsterdam: John Benjamins Publishing Company, vol. 38. [CrossRef]

Giannakidou, Anastasia. 1998. *Polarity Sensitivity as (Non)Veridical Dependency*. Amsterdam: John Benjamins Publishing Company.

Giannakidou, Anastasia. 2009. The Dependency of the Subjunctive Revisited: Temporal Semantics and Polarity. *Lingua* 119: 1883–908. [CrossRef]

Gil, Kook-Hee, and Heather Marsden. 2013. Existential Quantifiers in Second Language Acquisition. *Linguistic Approaches to Bilingualism* 3: 117–49. [CrossRef]

Green, David W., and Jubin Abutalebi. 2013. Language Control in Bilinguals: The Adaptive Control Hypothesis. *Journal of Cognitive Psychology* 25: 515–30. [CrossRef]

Grüter, Theres, Casey Lew-Williams, and Anne Fernald. 2012. Grammatical Gender in L2: A Production or a Real-Time Processing Problem? *Second Language Research* 28: 191–215. [CrossRef]

Guijarro-Fuentes, Pedro. 2012. The Acquisition of Interpretable Features in L2 Spanish: Personal A. *Bilingualism: Language and Cognition* 15: 701–20. [CrossRef]

Hopp, Holger. 2010. Ultimate Attainment in L2 Inflection: Performance Similarities between Non-Native and Native Speakers. *Lingua* 120: 901–31. [CrossRef]

Howard, Martin. 2008. Morpho-Syntactic Development in the Expression of Modality: The Subjunctive in French L2 Acquisition. *Canadian Journal of Applied Linguistics* 11: 171–92.

Hwang, Sun Hee, and Donna Lardiere. 2013. Plural-Marking in L2 Korean: A Feature-Based Approach. *Second Language Research* 29: 57–86. [CrossRef]

Ionin, Tania, and Eve Zyzik. 2014. Judgment and Interpretation Tasks in Second Language Research. *Annual Review of Applied Linguistics* 34: 37–64. [CrossRef]

Iverson, Michael, Paula Kempchinsky, and Jason Rothman. 2008. Interface Vulnerability and Knowledge of the Subjunctive/Indicative Distinction with Negated Epistemic Predicates in L2 Spanish. *Eurosla Yearbook* 8: 135–63. [CrossRef]

Kastronic, Laura. 2016. *A Comparative Variationist Approach to Morphosyntactic Variation in Hexagonal and Quebec French*. Ottawa: University of Ottawa.

Kato, Mary A., Sonia Cyrino, and Vilma Reche Corrêa. 2009. Brazilian Portuguese and the Recovery of Lost Clitics through Schooling. In *Minimalist Inquiries into Child and Adult Language Acquisition: Case Studies Across Portuguese*. Edited by Pires Acrisio and Jason Rothman. Berlin: Mouton De Gruyter, pp. 245–74.

Keating, Gregory D. 2009. Sensitivity to Violations of Gender Agreement in Native and Nonnative Spanish: An Eye-Movement Investigation. *Language Learning* 59: 503–35. [CrossRef]

Keating, Gregory D. 2010. The Effects of Linear Distance and Working Memory on the Processing of Gender Agreement in Spanish. In *Research in Second Language Processing and Parsing*. Edited by Bill VanPatten and Jill Jegerski. Amsterdam: John Benjamins Publishing Company, pp. 113–34. [CrossRef]

Keating, Gregory D., and Jill Jegerski. 2015. Experimental Designs in Sentence Processing Research. *Studies in Second Language Acquisition* 37: 1–32. [CrossRef]

Kempchinsky, Paula. 2009. What Can the Subjunctive Disjoint Reference Effect Tell Us about the Subjunctive? *Lingua* 119: 1788–810. [CrossRef]

Kroll, Judith F., and Ellen Bialystok. 2013. Understanding the Consequences of Bilingualism for Language Processing and Cognition. *Journal of Cognitive Psychology* 25: 497–514. [CrossRef]

Kuznetsova, Alexandra, Per B. Brockhoff, and Rune H. B. Christensen. 2017. LmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software* 82. [CrossRef]

Lardiere, Donna. 1998. Dissociating Syntax from Morphology in a Divergent L2 End-State Grammar. *Second Language Research* 14: 359–75. [CrossRef]

Lardiere, Donna. 2009. Some Thoughts on the Contrastive Analysis of Features in Second Language Acquisition. *Second Language Research* 25: 173–227. [CrossRef]

Lemhöfer, Kristin, and Mirjam Broersma. 2012. Introducing LexTALE: A Quick and Valid Lexical Test for Advanced Learners of English. *Behavior Research Methods* 44: 325–43. [CrossRef]

Lenth, Russell. 2018. *Emmeans: Estimated Marginal Means, Aka Leastsquares Means.* R package version 1.1.2. Available online: https://CRAN.R-project.org/package=emmeans (accessed on 13 February 2020).

Li, Ping, Fan Zhang, Anya Yu, and Xiaowei Zhao. 2019. Language History Questionnaire (LHQ3): An Enhanced Tool for Assessing Multilingual Experience. *Bilingualism: Language and Cognition*. [CrossRef]

Lim, Jung Hyun, and Kiel Christianson. 2015. Second Language Sensitivity to Agreement Errors: Evidence from Eye Movements during Comprehension and Translation. *Applied Psycholinguistics* 36: 1283–315. [CrossRef]

Liversedge, Simon P., and John M. Findlay. 2000. Saccadic Eye Movements and Cognition. *Trends in Cognitive Sciences* 4: 6–14. [CrossRef]

Marinis, Theodore, Leah Roberts, Claudia Felser, and Harald Clahsen. 2005. Gaps in Second Language Sentence Processing. *Studies in Second Language Acquisition* 27: 53–78. [CrossRef]

McDonald, Janet 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: 381–401. [CrossRef]

McManus, Kevin, and Rosamond Mitchell. 2015. Subjunctive Use and Development in L2 French: A Longitudinal Study. *Language, Interaction and Acquisition* 6: 42–73. [CrossRef]

Meisel, Jürgen M., Martin Elsig, and Matthias Bonnesen. 2011. Delayed Grammatical Acquisition in First Language Development. *Linguistic Approaches to Bilingualism* 1: 347–90. [CrossRef]

Mulder, Kimberley, and Jan H. Hulstijn. 2011. Linguistic Skills of Adult Native Speakers, as a Function of Age and Level of Education. *Applied Linguistics* 32: 475–94. [CrossRef]

O'Connor DiVito, N. 1997. *Patterns Across Spoken and Written French: Empirical Research on the Interaction among Forms, Functions and Genres*. Boston: Houghton Mifflin.

Osch, Brechje van, Aafke Hulk, Suzanne Aalberse, and Petra Sleeman. 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: 25. [CrossRef]

Pakulak, Eric, and Helen J. Neville. 2010. Proficiency Differences in Syntactic Processing of Monolingual Native Speakers Indexed by Event-Related Potentials. *Journal of Cognitive Neuroscience* 22: 2728–44. [CrossRef] [PubMed]

Papadopoulou, Despina, and Harald Clahsen. 2003. Parsing Strategies in L1 and L2 Sentence Processing. *Studies in Second Language Acquisition* 25: 501–28. [CrossRef]
Pérez-Leroux, Ana Teresa. 1998. The Acquisition of Mood Selection in Spanish Relative Clauses. *Journal of Child Language* 25: 585–604. [CrossRef] [PubMed]
Pires, Acrisio, and Jason Rothman. 2009. Acquisition of Brazilian Portuguese in Late Childhood: Implications for Syntactic Theory and Language Change. In *Minimalist Inquiries into Child and Adult Language Acquisition*. Edited by Pires Acrisio and Jason Rothman. Berlin: Mouton de Gruyter, pp. 129–54.
Plonsky, Luke, Emma Marsden, Dustin Crowther, Susan M Gass, and Patti Spinner. 2020. A Methodological Synthesis and Meta-Analysis of Judgment Tasks in Second Language Research. *Second Language Research* 36: 583–621. [CrossRef]
Poplack, Shana, Allison Lealess, and Nathalie Dion. 2013. The Evolving Grammar of the French Subjunctive. *Probus* 25: 139–95. [CrossRef]
Poplack, Shana. 1989. The Care and Handling of a Megacorpus: The Ottawa-Hull French Project. In *Language Change and Variation*. Edited by Fasold Ralph W. and Deborah Schiffrin. Amsterdam: John Benjamins Publishing Company, pp. 411–51.
Poplack, Shana. 1990. Prescription, Intuition et Usage: Le Subjonctif Français et La Variabilité Inhérente. *Langage et Société* 54: 5–33. [CrossRef]
Portner, Paul. 2018. *Mood*. Oxford: Oxford University Press.
Quer, Josep. 1998. *No Mood at the Interface*. The Hague: Holland Academic Graphics.
Quer, Josep. 2001. Interpreting Mood. *Probus* 13: 81–111. [CrossRef]
Quer, Josep. 2009. Twists of Mood: The Distribution and Interpretation of Indicative and Subjunctive. *Lingua* 119: 1779–87. [CrossRef]
R Development Core Team. 2014. *R: A Language and Environment for Statistical Computing*; Vienna: R Foundation for Statistical Computing. Available online: http://www.rproject.org (accessed on 27 January 2020).
Rayner, Keith, and Alexander Pollatsek. 1989. *The Psychology of Reading*. Englewood Cliffs: Prentice-Hall, Inc.
Rayner, Keith, Tessa Warren, Barbara J. Juhasz, and Simon P. Liversedge. 2004. The Effect of Plausibility on Eye Movements in Reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 30: 1290–301. [CrossRef] [PubMed]
Rayner, Keith. 1998. Eye Movements in Reading and Information Processing: 20 Years of Research. *Psychological Bulletin* 124: 372–422. [CrossRef] [PubMed]
Rayner, Keith. 2009. Eye Movements and Attention in Reading, Scene Perception, and Visual Search. *Quarterly Journal of Experimental Psychology* 62: 1457–506. [CrossRef] [PubMed]
Shimanskaya, Elena, and Roumyana Slabakova. 2017. Re-Assembling Objects: A New Look at the L2 Acquisition of Pronominal Clitics. *Bilingualism: Language and Cognition* 20: 512–29. [CrossRef]
Slabakova, Roumyana. 2015. The Effect of Construction Frequency and Native Transfer on Second Language Knowledge of the Syntax–Discourse Interface. *Applied Psycholinguistics* 36: 671–99. [CrossRef]
Slabakova, Roumyana. 2018. Inflectional Morphology. In *The Handbook of Advanced Proficiency in Second Language Acquisition*. Hoboken: Wiley, pp. 381–400. [CrossRef]
Sorace, Antonella. 2011. Pinning down the Concept of 'Interface' in Bilingualism. *Linguistic Approaches to Bilingualism* 1: 1–33. [CrossRef]
Soutet, Olivier. 2000. *Le Subjonctif*. Paris: Editions Ophrys.
Stepanov, Arthur, Sara Andreetta, Penka Stateva, Adam Zawiszewski, and Itziar Laka. 2019. Anomaly Detection in Processing of Complex Syntax by Early L2 Learners. *Second Language Research*. [CrossRef]
Suzuki, Yuichi. 2017. Validity of New Measures of Implicit Knowledge: Distinguishing Implicit Knowledge from Automatized Explicit Knowledge. *Applied Psycholinguistics* 38: 1229–61. [CrossRef]
Svenonius, Peter. 2019. Syntactic Features. In *Oxford Research Encyclopedia of Linguistics*. Edited by Mark Aronoff. Oxford: Oxford University Press, pp. 1–37. [CrossRef]
VanPatten, Bill, and Jill Jegerski. 2010. Second Language Processing and Parsing. In *Research in Second Language Processing and Parsing*. Edited by Bill VanPatten and Jill Jegerski. Amsterdam: John Benjamins Publishing Company, pp. 3–24. [CrossRef]
von Fintel, Kai. 2006. Modality and Language. In *Encyclopedia of Philosophy*, 2nd ed. Edited by Donald M. Borchert. Detroit: MacMillan Reference USA, Available online: http://web.mit.edu/fintel/fintel-2006-modality.pdf (accessed on 2 February 2020).
Whong, Melinda, Kook-Hee Gil, and Heather Marsden. 2014. Beyond Paradigm: The 'What' and the 'How' of Classroom Research. *Second Language Research* 30: 551–68. [CrossRef]
Zribi-Hertz, Anne. 2011. Pour Un Modèle Diglossique de Description Du Français: Quelques Implications Théoriques, Didactiques et Méthodologiques. *Journal of French Language Studies* 21: 231–56. [CrossRef]