Université de Montréal

Verbing and Nouning in French: Toward an ecologically valid approach to sentence processing

by Lauren Fromont

École d'orthophonie et d'audiologie, Faculté de médecine

Thesis presented in partial fulfillment of the requirements of the degree of

Doctor of Philosophy in Biomedical sciences option Speech language Pathology

Was evaluated by the following committee members:

Karine Marcotte, PhD (UdeM, president)

Phaedra Royle, PhD (UdeM, supervisor)

Karsten Steinhauer, PhD (McGill University, co-supervisor)

Boutheina Jemel, PhD (UdeM, internal member)

Colin Phillips, PhD (University of Maryland, external member)

Montreal, May 2019

Résumé

La présente thèse utilise la technique des potentiels évoqués afin d'étudier les méchanismes neurocognitifs qui sous-tendent la compréhension de la phrase. Plus particulièrement, cette recherche vise à clarifier l'interaction entre les processus syntaxiques et sémantiques chez les locuteurs natifs et les apprenants d'une deuxième langue (L2). Le modèle "syntaxe en premier" (Friederici, 2002, 2011) prédit que les catégories syntaxiques sont analysées de façon précoce: ce stade est reflété par la composante ELAN (*Early anterior negativity*, Négativité antérieure gauche), qui est induite par les erreurs de catégorie syntaxique. De plus, ces erreurs semblent empêcher l'apparition de la composante N400 qui reflète les processus lexico-sémantiques. Ce phénomène est défini comme le bloquage sémantique (Friederici et al., 1999). Cependant, la plupart des études qui observent la ELAN utilisent des protocoles expérimentaux problématiques dans lesquels les différences entre les contextes qui précèdent la cible pourraient être à l'origine de résultats fallacieux expliquant à la fois l'apparente "ELAN" et l'absence de N400 (Steinhauer & Drury, 2012).

La première étude rééevalue l'approche de la "syntaxe en premier" en adoptant un paradigme expériemental novateur en français qui introduit des erreurs de catégorie syntaxique et les anomalies de sémantique lexicale. Ce dessin expérimental équilibré contrôle à la fois le mot-cible (nom vs. verbe) et le contexte qui le précède. Les résultats récoltés auprès de locuteurs natifs du français québécois ont révélé un complexe N400-P600 en réponse à toutes les anomalies, en contradiction avec les prédictions du modèle de Friederici. Les effets additifs des manipulations syntaxique et sémantique sur la N400 suggèrent la détection d'une incohérence entre la racine du mot qui avait été prédite et la cible, d'une part, et l'activation lexico-sémantique, d'autre part. Les réponses individuelles se sont pas caractérisées par une dominance vers la N400 ou la P600: au contraire, une onde biphasique est présente chez la majorité des participants. Cette activation peut donc être considérée comme un index fiable des mécanismes qui sous-tendent le traitement des structures syntagmatiques.

La deuxième étude se concentre sur les même processus chez les apprenants tardifs du français L2. L'hypothèse de la convergence (Green, 2003 ; Steinhauer, 2014) prédit que les apprenants d'une L2, s'ils atteignent un niveau avancé, mettent en place des processus de traitement en ligne similaires aux locuteurs natifs. Cependant, il est difficile de considérer en même temps un grand nombre de facteurs qui se rapportent à leurs compétences linguistiques, à l'exposition à la L2 et à l'âge d'acquisition. Cette étude continue d'explorer les différences inter-individuelles en modélisant les données de potentiels-évoqués avec les Forêts aléatoires, qui ont révélé que le pourcentage d'explosition au français ansi que le niveau de langue sont les prédicteurs les plus fiables pour expliquer les réponses électrophysiologiques des participants. Plus ceux-ci sont élevés, plus l'amplitude des composantes N400 et P600 augmente, ce qui confirme en partie les prédictions faites par l'hypothèse de la convergence.

En conclusion, le modèle de la "syntaxe en premier" n'est pas viable et doit être remplacé. Nous suggérons un nouveau paradigme basé sur une approche prédictive, où les informations sémantiques et syntaxiques sont activées en parallèle dans un premier temps, puis intégrées *via* un recrutement de mécanismes contrôlés. Ces derniers sont modérés par les capacités inter-individuelles reflétées par l'exposition et la performance.

Mots-clés : Traitement de la phrase, Potentiels évoqués cognitifs, Catégories syntaxiques, Sémantique lexicale, Acquisition d'une langue seconde, Forêts aléatoires, ELAN, N400, P600

Abstract

The present thesis uses event-related potentials (ERPs) to investigate neurocognitve mechanisms underlying sentence comprehension. In particular, these two experiments seek to clarify the interplay between syntactic and semantic processes in native speakers and second language learners. Friederici's (2002, 2011) "syntax-first" model predicts that syntactic categories are analyzed at the earliest stages of speech perception reflected by the ELAN (Early left anterior negativity), reported for syntactic category violations. Further, syntactic category violations seem to prevent the appearance of N400s (linked to lexical-semantic processing), a phenomenon known as "semantic blocking" (Friederici et al., 1999). However, a review article by Steinhauer and Drury (2012) argued that most ELAN studies used flawed designs, where pre-target context differences may have caused ELAN-like artifacts as well as the absence of N400s.

The first study reevaluates syntax-first approaches to sentence processing by implementing a novel paradigm in French that included correct sentences, pure syntactic category violations, lexical-semantic anomalies, and combined anomalies. This balanced design systematically controlled for target word (noun vs. verb) and the context immediately preceding it. Group results from native speakers of Quebec French revealed an N400-P600 complex in response to all anomalous conditions, providing strong evidence against the syntax-first and semantic blocking hypotheses. Additive effects of syntactic category and lexical-semantic anomalies on the N400 may reflect a mismatch detection between a predicted word-stem and the actual target, in parallel with lexical-semantic retrieval. An interactive rather than additive effect on the P600 reveals that the same neurocognitive resources are recruited for syntactic and semantic integration. Analyses of individual data showed that participants did not rely on one single cognitive mechanism reflected by either the N400 or the P600 effect but on both, suggesting that the biphasic N400-P600 ERP wave can indeed be considered to be an index of phrase-structure violation processing in most individuals.

The second study investigates the underlying mechanisms of phrase-structure building in late second language learners of French. The convergence hypothesis (Green, 2003; Steinhauer, 2014) predicts that second language learners can achieve native-like online-processing with sufficient proficiency. However, considering together different factors that relate to proficiency, exposure, and age of acquisition has proven challenging. This study further explores individual data modeling using a Random Forests approach. It revealed that daily usage and proficiency are the most reliable predictors in explaining the ERP responses, with N400 and P600 effects getting larger as these variables increased, partly confirming and extending the convergence hypothesis.

This thesis demonstrates that the “syntax-first” model is not viable and should be replaced. A new account is suggested, based on predictive approaches, where semantic and syntactic information are first used in parallel to facilitate retrieval, and then controlled mechanisms are recruited to analyze sentences at the interface of syntax and semantics. Those mechanisms are mediated by inter-individual abilities reflected by language exposure and performance.

Keywords: Sentence processing, Event-related potentials, Syntactic categories, Lexical-semantics, Second language acquisition, Random forests, ELAN, N400, P600

Table of contents

List of tables

List of figures

List of abbreviations

AoA: age of acquisition
AoE: age of exposure
DC: direct current
EEG: electroencephalogram
ELAN: early left anterior negativity
ERP: event-related potential
IIR: infinite impulse response
(L)AN: (left) anterior negativity
MUC: memory — unification — control
RDI: response dominance index
ROI: region of interest
S&D2012: Steinhauer & Drury (2012)
UdeM: Université de Montréal

À mes parents

Acknowledgements

I would like to thank the wonderful individuals who have made this project possible and, perhaps more importantly, who have shared with me their knowledge, skills, and time. Every single moment I spent discussing, teaching, being taught, or simply sharing thoughts contributed to making me a better researcher.

My supervisors, Dr Phaedra Royle et Dr Karsten Steinhauer, have given me the opportunity to work on this project while I was still a bit of an outsider with no strong connections to the field. I owe them the very basis upon which I will continue building my knowledge and reflections. They provided me with invaluable training and, just as importantly, gave me the space I needed to learn novel techniques on my own as well. They trusted my decisions and challenged me when I needed to be. Thank you.

I would like to thank the research assistants and students who participated in this project: Audrey Borgus, Megan Honigman, Lisa Martignetti, Irina Perlitch, Mélanie Weiss, Alice Xue. This project would have not been possible at all without you. Thank you to Dr Royle's current and former lab members: Emilie Courteau, Dr Elena Kulinich, Dr Alexandra Marquis, Dr Marie Pourquié, Ariane Saint-Denis. Thank you to Dr Steinhauer's former lab members: Fayden Bokhari, Dr Kristina Kasparian, Dr Stefanie Nickels, Dr Efrat Pauker. It has been a pleasure to work alongside you.

I am extremely grateful to Dr Steinhauer's current students: Alexandre Herbay, Anastasia Glushko, and Max Wolpert—I have been so lucky to be your labmate. I truly look up to you three, and consider you among the smartest, kindest, and funniest individuals I have ever met. I am proud to have worked with you, I am thankful for every bit of support you have given me, and I hope I can do the same for you. You are my colleagues, my teachers, my inspirations, and above all, my friends.

Thank you to all the other wonderful persons that entered my life (both in and outside of Academia) and turned Montreal into a home I would not want to leave: Alberto, Baptiste, Gaston, Marcel, Marie, and Océane, as well as all the castellers of Montreal. A special thank you to Joel: I do not think many understand me as well as you do, and I am very grateful that you came into my life.

Thanks to all of you who, like me, advocate and fight for social justice in Academia: the Gender in Academia committee members at UdeM, Science and Policy exchange at McGill, the Gender committee at the Canadian Linguistic Association, and the Feminist Journal Club at UPF. I will take our discussions with me wherever I go, and hope for many more to come.

Finally, I want to thank my parents, to whom I dedicate this work. Merci de m'avoir soutenue toujours, de m'avoir fait confiance. Merci d'être les personnes merveilleuses que vous êtes. Vous n'avez pas idée à quel point je vous admire.

Funding acknowledgements

The research presented in this thesis received financial support from Social Sciences and Humanities Research Council (SSHRC grant # 435-2013-2052 - PI: Karsten Steinhauer, Phaedra Royle et al., Project on "Re-evaluating the temporal dynamics of syntactic online processing: Towards an ecologically valid model"). It was additionally supported by Canada Research Chair (CRC) program/CIHR (Project # 950-209843), Canada Foundation for Innovation (CFI/CRC project # 201876), Natural Sciences and Engineering Research Council (NSERC, grant # RGPIN 402678-11) (to PI: Karsten Steinhauer), SSHRC grant # 435-2015-1280 (PI : Phaedra Royle, Karsten Steinhauer et al), and FQRSC team grant # 2016-SE-188196 (PI: Lydia White, co-PI: Karsten Steinhauer, Phaedra Royle et al.).

Lauren Fromont was further supported by FQRSC's excellence scholarship program for foreign students (*Programme de bourses d'excellence pour étudiants étrangers - Fonds de recherche du Québec – Société et culture*), two graduate student stipends from the Centre for research on brain, language and music (CRBLM, Montréal), two excellence scholarships from the biomedical sciences program at the University of Montreal, and numerous travel awards from the CRBLM and the University of Montreal.

Author Contributions

Both manuscripts presented in this thesis contribute significantly and incrementally to the body of knowledge in neurolinguistics. Manuscript 1 presents a balanced design that allows for an evaluation of the syntax-first model, and brings an original cross-linguistic contribution to ERP language studies by focusing on Quebec French. Manuscript 2 provides evidence that further specifies the convergence hypothesis for second language learning (Steinhauer, 2014) and, to our knowledge, is the very first work that uses random forests to analyse ERP data. The manuscripts in this thesis were co-authored by Lauren Fromont (PhD candidate and first author), Phaedra Royle (supervisor and senior author on Manuscript 1), and Karsten Steinhauer (co-supervisor and senior author on Manuscript 2).

In the following, I provide a description of the contributions of the co-authors. Please note that while certain tasks were primarily carried out by a given author, every step resulted from a close collaboration between the three co-authors. The three authors equally contributed to data interpretation, and both co-authors gave their final approval on the manuscript versions included in this thesis and submitted to peer-review journals.

Lauren Fromont (PhD candidate): This author extended the experimental design, created and evaluated the stimuli materials, programmed the experiments, organized the recruitement of participants, scheduled participants and carried out the testing sessions with the help of research assistants (who are duely acknowledged in the manuscript submissions), trained research asssistants and supervised summer bursary students, pre-processed the data, carried out the statistical analyses (mixed-effects models and random forests), interpreted the results, and wrote the manuscripts.

Phaedra Royle (supervisor): This author provided the orginal idea of the experimental design, created stimuli sentences, provided the funds necessary to hire research assistants, provided the behavioral testing facilities for stimuli evaluation, gave extensive input and feedback during every step of the EEG experiment elaboration, analysis and interpretation, and edited the manuscripts.

Karsten Steinhauer (co-supervisor): This author provided the original hypothesis for reevaluating the syntax-first model (Steinhauer & Drury, 2012), pseudo-randomized the materials, provided the funds necessary to hire research assistants and recruit participants, provided the EEG testing facilities, gave extensive input and feedback during every step of the EEG experiment elaboration, analysis and interpetation, and edited the manuscripts.

General introduction

Figure 1. Calvin and Hobbes © Bill Watterson (January 25, 1993). Retrieved from: Homicidal Psycho Jungle Cat #9, reprinted with permission of Andrews McMeel Syndication. All rights reserved.

Why does verbing weird language? Syntactic categories (SCs), also called word categories, are functional information about words (e.g., noun, verb) that are used to build larger structures (for example, a determiner combines with a noun to form a determiner phrase, or DP). While it is possible in English to apply morphological derivation to words to change their syntactic category (e.g., "verb-*ing*", "weird-*s*"), unexpected categories tend to create processing difficulties. Writers rely on this to introduce comic (Figure 1), or poetic effects:

"He childed as I fathered"
(William Shakespeare, King Lear, Act III Scene 6)

In our everyday life, unexpected SCs are very productive and often memorable: Twitter expression "Don't @ me" is now so widely used that it even appears on T-shirts (retrieved from: https://www.teepublic.com/t-shirt/1645945-dont-at-me) and hip-hop music:

"I'm the hottest young'un, don't @ me"
(Jay Critch, Don't @ me, 2019. All rights reserved)

Under the assumption that SC information is instrumental in building hierachical structures, the processing of SCs during comprehension raises many important questions: How rapidly is this information used in order to build phrases? Are SCs special – is their processing distinct from lexical-semantics? And how do second language learners process this information?

In order to answer these questions, neurolinguistics online methods such as event-related brain potentials (ERPs) offer two major advantages that other techniques cannot. First, their excellent temporal resolution makes them a perfect tool to study the specific cognitive mechanisms that underlie different aspects of online sentence processing. Second, ERPs are a fined-grained multi-dimensional measure that allows researchers to draw qualitative differences (or similarities) about the underlying cognitive mechanisms of sentence processing in different populations, even when their behavior is indistinguishable (Steinhauer, 2014).

In past decades, ERP research has significantly contributed to the development of dynamic sentence processing models. Event-related potential components are characterized by their polarity (positive or negative), latency (in ms), and topographical distribution on the scalp. Twenty years ago, researchers seemed to have reached a clear picture about the underlying cognitive function of several language-related ERP components: lexical-semantic anomalies would elicit a larger N400 (Kutas & Hillyard, 1980), morphological agreement errors a (left) anterior negativity (L)AN) (Kutas & Hillyard, 1983), and ungrammatical sentences a P600 effect (Osterhout & Mobley, 1995). However, it has become evident that these components are not tied to specific linguistic processes (e.g., a P600 can be observed in response to semantic anomalies, Meerendonk, Kolk, Vissers, & Chwilla, 2008) and that most or all of them may not even be language-specific at all (e.g., Steinhauer, 2014). Further, it has been suggested that certain well-established language-related components may, at least in part, be artefacts resulting from flawed experimental manipulations (Steinhauer & Drury, 2012), problematic approaches to ERP analysis, and the process of averaging across overlapping components (Tanner, 2015). Moreover, individual ERP responses show that specific components such as the N400 and the P600 are not consistently observed across participants within the same experiment, thus prompting us to reevaluate their functional interpretation and to consider distinct psycholinguistic processing approaches even among native speakers. Neurolinguists therefore need to explicitely consider the interrelations between theoretical sentence-comprehension frameworks and methodological considerations linked to the ERP technique that they are using to evaluate those models.

This chapter will first present ERP components observed in language studies, and their functional interpretation. Next, three sentence processing models that propose distinct hypotheses on the dynamic processes underlying sentence processing will be discussed. Possible causes for apparent inconsistencies and variability across studies will be explored, with a particular focus on the interplay between SC identification and lexical-semantic anomalies. Finally, these issues will be discussed in the context of second language learning.

1. Language electrified: relevant ERP components

1.1. N400

The N400 is a centrally (sometimes broadly) distributed negative deflection that occurs between 300 and 500 ms after word presentation. N400 effects were famously discovered in a study comparing target words that were either congruent or incongruent with the sentence context ("He spread the warm bread with butter/?socks", Kutas & Hillyard, 1980). Since then, it has been observed in response to many experimental manipulations on lexical-semantic processing, with effects ranging from word frequency to the integration of world knowledge information (*Dutch trains are yellow/?white and very crowded*, Hagoort, Hald, Bastiaansen, & Petersson, 2004; see Kutas & Federmeier, 2011, for a review). There are two interpretations of the N400 effect: an "integration" and a "retrieval" view (Lau, Phillips, & Poeppel, 2008). The integration view states that the N400 reflects combinatorial processes (i.e. integration with the preceding context, e.g., Friederici, 2002; Baggio & Hagoort, 2011). Recently, however, most accounts have shifted towards a retrieval perspective of the N400 (e.g., Brouwer & Hoeks, 2013; Kuperberg, Caplan, Sitnikova, Eddy, & Holcomb, 2006; Lau et al., 2008). According to this view, the N400 effect is observed not because semantic anomalies induce a larger effect, but because a lower N400 amplitude is associated with easier lexical retrieval. In other words, the N400 is always present and reflects activation of lexical representation in the long-term memory (Lau et al., 2008). Several factors may occur simultaneously to facilitate lexical retrieval, and therefore modulate the N400: these range from "pure" priming leading to automatic-spreading activation (Kiefer, 2002), or context effects leading to contextual prediction (Kim, Oines, & Sikos, 2016; Van Petten & Luka, 2006). Despite great variability, most functional imaging studies and lesion studies seem to point to the posterior temporal cortex as a good candidate for storage and access of lexical information (Lau et al., 2008). Note that while most evidence seems to be consistent with a "retrieval" account, there is still some discussion as to whether later portions of the N400 component could be reflecting post-lexical integration (Steinhauer, Royle, Drury, & Fromont, 2017). The N400 effect is not limited to purely lexical-semantic features: it has also been observed in response to agreement errors (Molinaro, Barber, & Carreiras, 2011; *The clerk at the clothing boutique was/*were* *severely underpaid*, Tanner & Van Hell, 2014). Thus, based on more recent findings, the N400 may reflect more general predictive mechanisms operating on both word roots and inflectional (and derivational) morphemes, including syntax-driven dependencies.

The association of the N400 component with morpho-syntactic (as opposed to semantic) processing is a relatively new step; by contrast, the ERP components that will be presented next have been associated with syntactic processing since their discovery.

1.2. ELAN

The early left anterior negativity (ELAN) has been claimed to be a specific marker of syntactic category (SC) identification (Friederici, 2002). As it is elicited particularly early – between 100-300 ms after word presentation or onset (and sometimes earlier) – the ELAN is taken to reflect rapid and automatic processes independent of task and experimental context (Hahne & Friederici, 1999, 2002). The experimental paradigm in seminal ERP studies (Friederici, Pfeifer, & Hahne, 1993; Friederici, Steinhauer, & Frisch, 1999) and its adaptation to other languages (Hinojosa, Martin-Loeches, Casado, Muñoz, & Rubia, 2003; Isel, Hahne, Maess, & Friederici, 2007; Rossi, Gugler, Friederici, & Hahne, 2006) relies on context manipulation. As illustrated in example sentences (1-2), this type of design keeps the target constant; the SC violation is created by inserting the preposition – determiner contraction *am* – ‘on-the’, which selects for a noun, before the verb target:

(1) *Die Bluse wurde <u>gebügelt</u>*. ‘The blouse was ironed.’ (Correct)
(2) *Die Bluse wurde am *<u>gebügelt</u>*. ‘The blouse was on-the[1] ironed.’ (SC violation)

[1] The hyphen in “on-the” emphasizes the contraction between the preposition and the determiner in the German word *am* (which can also translate in English as to-the, roughly equivalent to the French *au* = *à le*).

The ELAN effect appears very early because it is putatively triggered by the prefix *ge-* (that indexes most German past participles and is claimed by Friederici and colleagues to be a reliable verb marker). In a comprehensive review of experimental ELAN studies, Steinhauer and Drury (2012, henceforth S&D2012) develop two main arguments concerning this context manipulation paradigm: (i) it does *not* create outright syntactic violations at verb onset (prefix *ge-*), and (ii) it leads to context effects that have consistently been misinterpreted as ELANs. Regarding (i), S&D2012 point out that in German, like in English, most past participles can be used as pre-nominal adjectival modifiers such as in the perfectly grammatical sentence (3):

(3) *Die Bluse wurde am gebügelten Jackett mit Nadeln befestigt*.
The blouse was to-the ironed jacket with pins attached.
'The blouse was attached to the ironed jacket with pins'

Thus, the occurrence of an ELAN in this type of sentence cannot be associated with a failure in phrase-structure building due to a SC violation, because there is no violation when the prefix *ge-* is heard. Moreover, since many nouns and adjectives in German start with *ge-*, S&D2012 argue that the syntactic category identification cannot be derived through a strictly bottom-up process based on phonetic information at word onset. They rather argue that the early activation reflected by an ELAN could be explained by phonotactic strategies developed by participants due to a limited experimental design: that is, participants learn during the experimental session that prepositions followed by a word starting in *ge-* will always result in a syntax error (see also Hagoort, 2003b for a similar interpretation).

There are two types of context effects (ii) that may affect the ERP components and be misinterpreted as early ELANs. First, the *spillover effect* is a transient effect observed when ERP differences (e.g., P600 effects) triggered by a context word preceding the target only occur (or continue to gain amplitude) *after* target word onset. In this scenario, a pre-target baseline correction would not compensate for the delayed context effect, and the resulting ERP difference is likely to be misinterpreted as an ERP effect triggered by the target word. Second, the *DC offset effect* describes sustained waves created by an artifactual vertical shift of an entire waveform (towards more positive or more negative values) due to baseline correction. This can happen when the words preceding the target in the control and violation conditions elicit different effects that show up within the pre-target baseline interval. For example, a content word (e.g., a verb) typically elicits a larger N400 than a function word (e.g., a preposition) (Frank, Otten, Galli, & Vigliocco, 2015). Yet, in many studies the standard N400 interval (300-500 ms *post* word onset) coincides with the 200 ms *pre*-stimulus interval for the *next* word that serves as its standard baseline interval. As baseline corrections attempt to compensate (i.e., remove) pre-stimulus differences between conditions by forcing their waveforms together, the waveform with the larger (pre-target) content word-N400 would be moved towards more positive amplitudes, and the other condition (with a pre-target function word) towards more negative amplitudes, resulting in measurable early negativity right after target word onset. Regrettably, this scenario is typical for many SC violation studies where – importantly – the function word (e.g., a preposition) appeared in the SC violation condition (and a content word in the control), thereby creating early sustained negativities after target word onsets that have been mistaken for ELAN components elicited by the target word (see S&D2012 for details). Given that a large number of ELAN findings may be subject to methodological issues, it remains unclear to what extent ELANs are the result of flawed designs or reflexes of genuine neurocognitive processes linked to SC violations.

While ELANs have been taken to reflect SC identification processes, the LAN component, similarly distributed but elicited at later latencies, has been interpreted as a more general marker of morpho-syntactic processing.

1.3. (L)AN

The (left) anterior negativity, or (L)AN, is observed around 300–500 ms, usually at left anterior sites even though its topography varies: it can be largest at left temporal sites (e.g., T5, Newman, Ullman, Pancheva, Waligura, & Neville, 2007; Steinhauer et al., 2010) in the visual modality, or have an anterior – but not so lateralized – distribution in the auditory modality (e.g., Silva-Pereyra, Rivera-Gaxiola, & Kuhl, 2005). LAN effects are mostly – but not consistently (Molinaro et al., 2011) – elicited in response to agreement violations. Common cases of agreement errors that elicit a (L)AN include: subject-verb agreement (e.g., *The officials *hope**s** to succeed*, Osterhout & Mobley, 1995), number agreement on the noun (e.g., *All turtles have four *legØ and a tail*, Kutas & Hillyard, 1983), and gender agreement (e.g., determiner-noun and adjective-noun in Spanish, Barber & Carreiras, 2005).

LAN effects have also been observed for SC violations (e.g., *Tengo que corer muchas millas / *muchas millas correr*[2] Bowden, Steinhauer, Sanz, & Ullman, 2013; Neville, Nicol, Barss, Forster, & Garrett, 1991; Steinhauer, White, & Drury, 2009). Since S&D2012 cast serious doubt on the validity of the ELAN as a real ERP component, syntactic category identification has tended to be associated with E/LANs, without further discussion about theoretical implications of latency differences between the two components. For example, in a study of SC violations in French, Brusini et al. (2017) define LAN latencies as "typically

---

[2] I have to run many miles / *many miles run.

appear[ing] between 100 and 400 ms". While this seems to imply that both ELAN and LAN effects reflect the same cognitive processes with different latencies, serial approaches to sentence processing (and in particular, Friederici, 2002, 2011, 2012) clearly define these two components as functionally dissociable.

As mentioned in §1.1. of this introduction, N400 effects have also been observed in response to some agreement errors, leading to continuing discussions about the functional relationship between LANs and N400s. On the one hand, LANs and N400s are interpreted as two extremes on one continuum: agreement errors that rely more strongly on morphological inflexion (e.g., *The elected officials *hopes to succeed*, Osterhout & Mobley, 1995) would elicit a LAN while word-stem mismatches (e.g., *The clerk at the clothing boutique was/*were severely underpaid*, Tanner & Van Hell, 2014) would trigger an N400 (Molinaro, Barber, Caffarra, & Carreiras, 2015). On the other hand, it has been suggested that the LAN is an N400 "in disguise". It only *appears* to be anterior and left lateralized because it overlaps with the P600 effect: since the latter is more posterior and right lateralized, it cancels the N400 out, except at left-anterior electrodes (Tanner, 2015). However, the presence of a LAN in the absence of a P600 effect (e.g., Mancini, Molinaro, Rizzi, & Carreiras, 2011) suggests that at least *some* left anterior negativities are genuine.

The next section focuses on the P600 component, which has been observed in response to (morpho-)syntactic violations, sometimes in conjunction with a LAN effect, and other times on its own.

1.4. P600

The P600 effect is a long-lasting (up to 1000 ms) and posterior positivity that starts around 500 ms, and has been observed in response to a range of syntactic manipulations.

These include local dependencies: SC violations (Friederici, Gunter, Hahne, & Mauth, 2004a; Friederici et al., 1993, 1999), agreement errors (Molinaro et al., 2011; Osterhout & Mobley, 1995), syntactic ambiguities (Carreiras, Salillas, & Barber, 2004), garden-path sentences (*The lawyer charged the defendant was lying*, Osterhout, Holcomb, & Swinney, 1994), and long-distance dependencies (Phillips, Kazanina, & Abada, 2005). It has also been observed in response to thematic role reversals – effect known as the "Semantic P600" (*For breakfast the eggs would eat toast*, Kuperberg et al., 2006), and lexical-semantic anomalies (Meerendonk et al., 2008). To account for all of these phenomena, interpretations of the P600 are manifold. It is often suggested that it reflects integration (as opposed to retrieval, Brouwer & Hoeks, 2013), or (somewhat underspecified) combinatorial processes (Osterhout, Mclaughlin, Pitk, Frenck-Mestre, & Molinaro, 2006; Tanner & Van Hell, 2014). Friederici (2002) suggests that the P600 effect reflects controlled revision and repair processes. Hagoort (2003a) interprets it as an index of unification and selection among competing structural analyses. The monitoring theory (Meerendonk et al., 2008; Vissers, Kolk, van de Meerendonk, & Chwilla, 2008) suggests that the P600 is elicited when a strong conflict between what is expected and what is observed triggers reanalysis. This range of P600 accounts may be best understood as suggesting that "it is not a monolithic component" (Friederici, Mecklinger, Spencer, Steinhauer, & Donchin, 2001).

Importantly, the P600 is task-dependent (Hahne & Friederici, 2002; Hasting & Kotz, 2008; Royle, Drury, & Steinhauer, 2013) and has been found to be larger when a sentence is classified as unacceptable within specific experimental contexts, e.g., where many sentences are grammatical. In this respect, it bears similarities with the P3b component, which has been observed in attended oddball paradigms where the deviant elicits a large parietal P300 that increases in amplitude the rarer and the more task-relevant the deviant is. The "P600 as P300 hypothesis" (Coulson, King, & Kutas, 1998; Sassenhagen, Schlesewsky, & Bornkessel-Schlesewsky, 2014) therefore interprets the P600 as a reflection of working memory updating. This view has been independently supported by findings that verbal working memory abilities correlate with the P600 effect (Kim, Oines, & Miyake, 2018; Vos, Gunter, Kolk, & Mulder, 2001).

The P3b sometimes appears with another ERP component also related to task processing and attention: the P3a.

1.5. P3a/Frontal P600

The P3a is a positive wave that appears at frontal electrodes, generally around 300–500 ms after stimulus onset (although its latency range varies). It is a domain-general component that reflects surprisal and re-allocation of attention in response to an oddball (Squires, Squires, & Hillyard, 1975; see Polich, 2007, for a review). However, P3a-like effects have also been observed in language studies and are sometimes referred to as "early P600s" (Molinaro et al., 2011), or "frontal P600s" (Kaan & Swaab, 2003). They are normally visible right after the LAN or N400 effect — around 500 ms. Such frontal positivities have often been described as being language- (and even syntax-) specific. They have been attributed to multiple processes including (i) difficulties in integrating a constituent with a previous context related or to non-preferred sentence continuations (as opposed to ungrammatical ones, Hagoort, Brown, & Osterhout, 1999), (ii) ambiguous structures (Kaan & Swaab, 2003), or (iii) agreement error processing (Barber & Carreiras, 2005; Molinaro, Vespignani, Zamparelli, & Job, 2011) and (iv) number incongruency processing with grammatical input (Courteau, Martignetti, Steinhauer, & Royle, 2019). However, none of these interpretations really accounts for the many instances in which frontal positivities are observed; in fact, these are not restricted to syntactic processing as they have been observed in response to logical semantic anomalies (Bokhari, 2015). Alternatively, Kasparian, Vespignani, and Steinhauer (2017) suggest that the early positivity could be interpreted as a P3a in experiments where anomalies are less predictable: for example, when they occur early in the sentence without much context, or when they occur sporadically in the experiment.

Sentence processing models adopt different views on how ERP components in language studies are interrelated, and on their functional interpretation. We present three models that are relevant to the present study in the next section.

2. Approaches to sentence processing in neurolinguistics

Theories of language comprehension aim to clarify the mechanisms that are at play from the moment a word is presented in a sentence context until the overall message of this sentence is integrated. Specifically, we seek to understand whether these mechanisms subserve distinct linguistic functions (e.g., syntax vs. semantics) or general ones that can either be language-specific (e.g. retrieval in verbal memory), or not (e.g. prediction error). In addition to characterizing the nature of these mechanisms, we want to specify their temporal organization: whether they take place serially or in parallel, and whether they are encapsulated or interact with one another.

2.1. A serial approach: Friederici's "syntax-first" model

Serial, "syntax-first" approaches adopt the general principle of a modular organisation of the syntactic and semantic processes that underlie sentence comprehension. Further, they posit that syntactic analysis of local relationships between constituents prevails over lexical-semantic integration: while multiple candidates may be activated during lexical access, the parser rapidly commits to one syntactic analysis of the constituent structure (and revises it if necessary). According to Frazier's Minimal attachment theory (1987), the rapid analysis of local relationships between constituents is governed by the relative simplicity of syntactic structures: the listener chooses the simplest one, independent of frequency of use and lexical-semantic characteristics.

The "syntax-first" model by Friederici (2002, 2011, 2012) is most inspired by Frazier (1987)**.** This serial and modular model posits an encapsulated "syntax module" in the brain. Sentence processing mechanisms are divided into three phases, which are illustrated in Figure 2, along with the ERP components that reflect them.

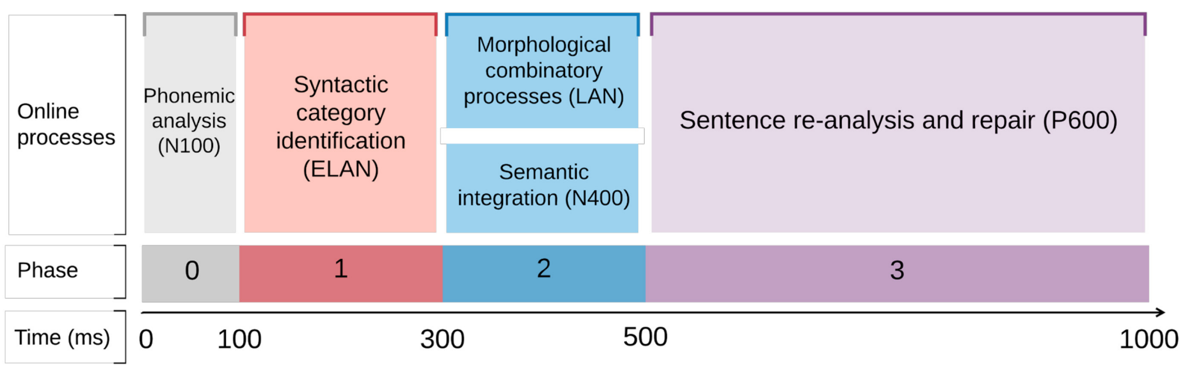


Figure 2. The "syntax-first" model of sentence comprehension. Adapted from Friederici (2002, 2011).

For each incoming word, SC information is retrieved first (Phase 1): the parser is "blind" to any other types of information (e.g., word meaning or agreement features). If SC identification leads to a felicitous phrase-structure representation, lexical-semantic information (reflected by the N400) and morpho-syntactic information (reflected by the LAN) may be integrated in parallel during Phase 2. Finally, in Phase 3, structural repair processes take place in response to syntactic violations or difficulties.

In the presence of an SC violation, the lack of a grammatical phrase-structure representation prevents morpho-syntactic feature checking: agreement can only occur within intact phrase markers. That is, participants presented with an anomalous sentence that combines an SC error and an agreement error (e.g., *The boy in-the ø *sing a song*[3], **Rossi et al 2005**) elicit an ELAN *but no LAN*. Likewise, semantic-thematic interpretation of the target word (**Frisch et al, 2004**) and lexical-semantic integration are blocked – thus suppressing the N400 effect (Friederici et al, 1999: "Semantic blocking" is more thoroughly discussed in Chapters 2 and 3).

These findings are important because they provide evidence that phrase-structure building is not only fast and automatic, but also encapsulated, and conditions later stages of sentence processing – note that semantic integration and agreement processing occur in parallel, independently from each other (Gunter, Friederici, & Schriefers, 2000). Taken at face value, these data suggest that rapid SC identification cannot simply reflect the early *availability* of the SC information relative to other cues, but rather the speed at which this information is being *used*. Indeed, in the Rossi et al's (2005) experiment, SC information and agreement marking are made available at the same time, yet the parser focuses on the SC cues first. One yet more compelling piece of evidence is that the "semantic blocking" effect can be replicated *even when* lexical-semantic information is presented first (i.e., the word stem preceding the suffix in the auditory domain): the parser waits until the SC information is made

[3] 'Der Junge im ø *singst ein Lied'

available in the word suffix (e.g., *The bush was* *replanted* *(...)' / 'The bush was despite* **replanted* *(...)*[4], Friederici, Hahne, Gunter & Mauth, 2004). While one can hypothesise that, in these experiments in German, SC information can be rapidly retrieved by stripping affixes, it remains unclear how the parser does this in languages where there is no clear affixation. Nevertheless, similar evidence has been provided in French (Isel et al., 2007), showing that combined SC and lexical-semantic anomalies elicit an ELAN but no N400 in words where no morphosyntactic cue can serve as an SC marker (e.g., *dort* 'sleep'). That is, the very nature of lexical access during sentence comprehension is underspecified. The influential extended Argument Dependency Model (Bornkessel & Schlesewsky, 2006), also adopted the idea that ELANs exclusively reflect early identification of SCs, but do specify that this process is morphologically driven.

By contrast, parallel, interactive accounts of sentence processing generally consider that structural representations are encoded in the lexicon, and that, therefore, lexical and syntactic information are processed in parallel and characterized by somewhat similar mechanisms that have the ability to influence one another.

2.2. The parallel approach

According to Hagoort's Memory — Unification — Control model (MUC, Hagoort, 2005, 2016, inspired by Vosse & Kempen, 2000), phonological, syntactic, and semantic information are processed in parallel though highly interactive streams. We will describe the semantic and syntactic processes that take place when a word like "reader" is encountered in example sentence (4).

---

[4] 'Der Strauch wurde verpflanzt (…) / Der Strauch wurde trotz *verpflanzt (…)'

(4) This sentence annoyed the reader.

The lexical representation of "reader" is activated from Memory: it contains all its possible meanings (that will be used for semantic Unification) and all its possible syntactic frames (that will serve for syntactic Unification). Syntactic frames are hierarchical structural environments associated with a given lexical item (Jackendoff, 2002): for "reader" it may include its projection as an NP (the root node), its function as a head (the functional node), and its category *per se* as a noun (the foot node). Once this representation is activated from Memory, Unification consists in a selection mechanism that "chooses" the right frame for it: for example, after Unification of "the + reader", the root node of the NP "reader" is the foot node and complement of "the" (DP—head—Determiner; similar to Head-driven phrase structure grammar, Pollard & Sag, 1994).

Semantic Unification is a constructive process that results in a discourse model. Although, in the present work, integration and unification are used interchangeably (as opposed to retrieval), note that Hagoort, Baggio, and Willems (2009) make a distinction between integration (different sources converge on a common memory representation) and Unification (a new representation is constructed that is not already available in Memory).

Semantic Unification (reflected by the N400 effect) and syntactic Unification (reflected by a LAN-P600 complex) are distinct processes in this model. However, they should be able to interact at any time: hypothetically, an SC error could elicit an N400 instead of a LAN if the SC resulted mostly in a violation of lexical-semantic expectations for this specific target. On the other hand, a LAN and no N400 could be observed in cases where the syntactic category violation is implemented via morphological affixes, and both correct and incorrect target words are either equally unlikely or equally likely to occur (i.e. if they have the same cloze probability: *The lumberjack dodged the vain propeller / *propelled on Tuesday*[5]. Hagoort, 2003b). In the case where a target in anomalous because of a conflict between syntactic and semantic information (e.g., in argument-structure violations ‘The hearty meal was devouring….’, (Kim & Osterhout, 2005), this conflict is resolved by obeying a “loser takes all” principle (Hagoort et al., 2009). In the aforementioned example, semantic associations are more (i) strongly constrained, because of the association between ‘meal’ and ‘devour’ and (ii) are taken into account earlier than syntactic ones. Since syntactic constraints are presumably weaker, the anomaly is perceived as a morphosyntactic violation (e.g., *devouring* instead of *devoured*), so the effect appears on the P600 (thus the weaker constraints “lose” and take on the ERP effect). We will return to this interpretation in the general discussion of the present thesis.

This model accommodates cases where one type of information is taken into account earlier than the other. From this perspective, ELANs are interpreted as LANs that *appear* early when a prefix unambiguously points to the wrong syntactic category (Hagoort, 2003b). When the information is provided by the suffix of the word, a LAN is observed, and lexical-semantic processing is not blocked (Van den Brink & Hagoort, 2004).

Serial and parallel approaches to sentence processing adopt very different hypotheses on how syntactic and semantic information is processed. Nevertheless, both paradigms rely on a strong distinction between mechanisms underlying the processing of these two types of

---

[5] ‘De houthakker ontweek de ijdele schroef/ *schroeft op dinsdag’

linguistic cues. More recently, sentence processing accounts have blurred the distinction between syntax and semantics.

2.3. The predictive approach

We have explained that one limitation with Friederici's model is that it does not specify how SC information can be activated without consulting the lexicon in the absence of affixation. Further, it does not specify whether this early effect reflects hard-wired syntactic preferences that make the analysis rapid and straightforward (Frazier, 1987; Goucha, Zaccarella, & Friederici, 2017), or at least partly reflects predictive processes that may be driven by contextual restrictions (Gibson, 1998; see also Kimball, 1975, for an early integration of predictive analyzers into syntax-first models), or the relative frequency of certain syntactic structures (see MacDonald, Pearlmutter, & Seidenberg, 1994, for a discussion). The idea that prediction guides parsing is therefore present in most serial models of sentence processing, except Frazier's, which excludes semantic – but not syntactic – prediction, and, seemingly, Friederici's – although her data can be interpreted using prediction (Lau, Stroud, Plesch & Phillips, 2006).

The predictive approach has gained increasing popularity over the past decade (one of six ERP presentations at the *CUNY 2018 human sentence processing* conference focusing on ERPs was explicitly about prediction). There are several ways this approach can be defined, but the present thesis (mostly) relies on work by Ellen Lau and collaborators (2006, 2008; but see also (Chow, Momma, Smith, Lau, & Phillips, 2016), and to some extent, on discussions between Darren Tanner and Nicola Molinaro (Molinaro et al., 2015; Tanner, 2015). Although this approach seeks to explain effects of predictive processes during sentence comprehension, relatively few studies have investigated —or observed— the effects of prediction *when* it happens (Molinaro, Giannelli, Caffarra, & Martin, 2017; Molinaro, Monsalve, & Lizarazu, 2016). Rather, researchers measure the ease of retrieval (quantified by the amplitude of the N400 effect) on the target word after it has already been predicted (or pre-activated). Two notable studies show that early activation of a target word may due to predictive processes that impact speed of retrieval. DeLong, Urbach, and Kutas (2005) manipulated the cloze probability of target nouns that, importantly, had different determiners (in sentence [5], high cloze probability "*a* kite" vs. low cloze probability "*an* airplane"). An N400 effect was observed on determiners that preceded low cloze nouns, suggesting target pre-activation through predictive processes. Note, however, that a large-scale study failed to replicate their results (Nieuwland et al., 2018).

(5) The day was breezy, so the boy went outside to fly <u>a kite</u> / <u>an airplane</u>

Furthermore, in the context of SC identification, Lau, Stroud, Plesch, & Phillips (2006) suggested that ELANs appeared in highly predictive contexts where the analysis precedes the input. They evaluated sensitivity of the ELAN to predictability by comparing ungrammatical sentences in a low constrained context (6) or a very predictive context (7). Indeed, the sentence in (6) does not necessarily predict an overt noun at the end of the possessor (the sentence: "Although Erica kissed Mary's mother, she did not kiss Dana's" is perfectly grammatical), while (7) strongly predicts a noun after Dana's. The authors hypothesized that an SC occurring in a less predictive context would reduce the ELAN effect.

(6) Although Erica kissed Mary's mother, she did not kiss Dana's *<u>of</u> the bride.
(7) Although the bridesmaid kissed Mary, she did not kiss Dana's *<u>of</u> the bride.

The most predictive context (7) elicited a larger ELAN than (6) at the target *of*, suggesting a possible alternative interpretation of the ELAN: it is observed in specific situations where predictions strongly constrain the set of possible candidates. However, this study is not exempt of possible methodological issues. First, the authors used an average reference, but still included anteriority and hemisphere as factors. Using an average reference promotes interactions between ERP effects and anteriority or hemisphere, since this type of averaging forces some parts of the scalp potentials to be negative and others to be positive, in order for the overall scalp average to be zero. Second, technically, this is a context manipulation design, because the possessive “Dana’s” in (6) may activate the covert noun “mother”. Analyzing the ERPs at the preposition onset, with a 100 ms pre-stimulus baseline, like the authors did, might lead to similar artefacts as the ones described by S&D2012. Kaan and collaborators (2016) used the same materials as Lau et al. (2006) but measured the ERPs at the onset of the possessive (e.g., “Dana’s”) and applied an averaged mastoids reference, a more standard approach for this type of stimulus. They did not replicate Lau et al.’s ELAN.

Therefore, the ELAN may not be a very good candidate to investigate the effects of prediction. However, the N400 has shown much more consistent results. Within the predictive framework, the N400 reflects the activation of a neighborhood of concepts and words through bottom-up (i.e., contextual) input. In a highly predictive context, if a speaker has been building sentence structure and meaning incrementally, then irrelevant competitors are rapidly suppressed. Manipulating structural predictability while keeping lexical-semantic priming constant has shown that some processes are rapid and incremental (e.g., negation relations, Nieuwland & Kuperberg, 2008) and others are slower (e.g., argument-structure relations, W. Y. Chow, Lau, Wang, & Phillips, 2018).

In order to make predictions, speakers need to make use of various linguistic cues delivered by the input. In particular, SCs may be identified using three types of criteria, which vary across languages (Basciano, 2017): (i) syntactic criteria that are based on the distribution of words and their compatibility with other SCs, (ii) morphological criteria that concern derivation, and (iii) lexical-semantic information criteria. As SC violations appear to result in distinct patterns across experiments in different languages, cross-linguistic differences have been proposed as an explanation for these apparent discrepancies.

3. Variability affects syntactic category identification

3.1. Cross-linguistic variability I: evidence from Mandarin Chinese

An explicit cross-linguistic approach was adopted in three ERP studies of SC processing in Mandarin Chinese. The following studies investigated whether the lack of morpho-syntactic markers to identify SCs would affect the time-course of phrase-structure building in Mandarin. In all three studies, an N400 effect was observed in response to SC violations that were either combined with lexical-semantic anomalies or not (see Chapter 2 §1). In particular, Yang, Wu, and Zhou, (2015) offer an interesting cross-linguistic perspective, since they use the Mandarin BEI structure (8) that closely parallels the passive used in the German studies from Friederici's lab:

(8) That piece of glass BEI is carefully wiped / *dishcloth […]. (English translation)

Yang and collaborators conclude that the presence of an N400 instead of an ELAN indicates that lexical-semantic processing has "primacy" in Chinese. However, this interpretation may not be quite reflective of the actual linguistic characteristics of Mandarin Chinese: as in most languages, it is not possible to rely solely on lexical-semantic information to identify SCs, so speakers have to rely on the distribution of words in specific contexts in order to do so (Tuijl, 2017).

The conflicting results between the German and the Mandarin Chinese studies may not be attributable to cross-linguistic differences alone. Indeed, one major methodological difference between the ERP paradigms used in Mandarin and in German (although not mentioned by the authors) is that the BEI paradigm exemplified in (8) relies on a target manipulation approach. It therefore avoids any kind of context-related artefacts that may explain ELANs and the absence of N400s (S&D2012). In the Mandarin studies, contexts are kept constant, but nouns are always associated with the SC violation. This is most probably due to the fact that SCs in Mandarin Chinese are generally described as embedded within one another: all verbs can be used as nouns, but not all nouns can be verbed. This puts a constraint on the experimental design, as one cannot create an SC violation on a verb. Unfortunately, such a target manipulation design makes it difficult to tease apart SC violation effects from target-specific lexical effects.

In the next section, we turn to studies that focus on French SCs, as this language will be the focus of our experiments.

3.2. Cross-linguistic variability II: evidence from French

In French, syntactic categories are much more encapsulated than in Mandarin Chinese and may be identified by derivational morphology. For example, verbs are categorized in three groups depending on their suffix in the infinitive form: *-er*, *-ir*, or *-re*. However, inflections are sometimes opaque or irregular (e.g., *je bois* "I drink" and *elle boit* "she drinks" are both pronounced [bwa]). Word-stems in French are also reliable SC markers, even though homonyms exist, as in most languages (e.g., *du bois* [bwa] is a noun that means "wood").

The first studies focusing on French SCs used a context manipulation approach similar to Friederici's paradigm (Isel, 2007; Isel et al., 2007) in the auditory modality. Syntactic category errors (9-10) were created by removing the noun "house" from correct sentences so the verb "is sleeping" would directly follow the determiner of the prepositional phrase:

(9) *L'enfant qui est dans la maison dort*. 'The child who is in the house is sleeping.'
(10) *Le chauffeur qui est dans la *dort*. 'The driver[6] who is in the *is sleeping.'

As SC information was not provided by a prefix (and never is in French), the "ELAN" observed by Isel and colleagues at 150 ms is even less believable in that it actually appears in the ERP before the SC violation is present in the speech signal (i.e. within the stem or at word's end). In addition, although a central, N400-like, negativity is visible in the ERP plots for syntactic and semantic anomalies alike, the authors do not even consider the possibility of an N400 effect (that would contradict the semantic blocking hypothesis), and instead interpret the whole negative waveform as an ELAN effect. Finally, they state that similar results were obtained regardless of the baseline (i.e. pre- or post-stimulus onset) without providing the data supporting this claim.

A set of three auditory ERP studies of French (Bernal, Dehaene-Lambertz, Millotte, & Christophe, 2010; Brusini, Dehaene-Lambertz, Dutat, Goffinet, & Christophe, 2016; Brusini et al., 2017) investigated whether toddlers (18 and 24 months old) compute syntactic structure

[6] Subjects are different because the same participant heard both sentences, which means that target words were repeated several times.

online, or rely on local relationships to comprehend sentences. They created a design in the auditory modality where a French function word immediately preceding the target word (noun or verb) is ambiguous (i.e. *la* ‘the$_{FEM}$/it$_{FEM}$/her’). *La* can be either an object clitic pronoun such as “it/her” that precedes a verb phrase, or a definite determiner – “the”, which precedes a noun phrase. The target word could be either correct (11) or a SC violation in the sentence context (12). All sentences were embedded in short stories (approximately five sentences each), which contained both correct and incorrect target words.

(11) *Elle donne la fraise*. ‘She gives the strawberry’
(12) *Alors elle la *fraise sans y penser.* ‘Then she *strawberries it[7] without thinking’

Adult-based controls showed negativities starting around 250 ms after target onset, followed by a P600 at around 550 ms. Infants elicited varying results across studies, sometimes eliciting an ELAN (Brusini et al., 2017), sometimes not (Brusini et al., 2016). If genuine, the early negativities can be attributed to the small set of target words ($N = 4$) used in the experiment: participants could rapidly guess by the first syllable what the target word would be. However, it is difficult to determine whether these effects are genuine or artefacts due to design flaws. First, the target *fraise* can also be a transitive verb in French: *elle la fraise* is therefore grammatical and means “she mills it”. It is not very frequent, but adult speakers would certainly know this word: the authors created a lexical-semantic incongruency at best. Second, and more generally, there was little control over prosodic factors, the stimuli were presented in unbalanced contexts, and no control over phonemic (e.g., syllable structure) or lexical characteristics (e.g., frequency) was undertaken.

[7] The object pronoun *la* precedes the verb in French. See more details from our paradigm below.

In sum, there is quite some confusion around what processes actually subserve SC identification in the literature: many results are inconsistent, even in cases when the focus is on similar languages and structures (e.g., in French). It appears that at least some of these inconsistencies may be attributable to inter-individual variability.

3.3. Inter-individual variability: the "tradeoff" hypothesis

Apart from the "syntax-first" model, where the ELAN is *the* only component reflecting early processes, sentence processing models generally distinguish between "earlier processes" that are indexed in ERPs by negativities, and "later processes" reflected by the P600. While the presence or absence of either of these ERPs may vary as a function of the language or the experimental design, qualitative differences have also been observed across individuals for the same stimuli and in the same experiment. Specifically, some participants show a larger N400 and others a larger P600 effect in response to difficult words in garden-path sentences (Bornkessel, Fiebach, & Friederici, 2004), agreement errors (Tanner & Van Hell, 2014), lexical-semantic incongruencies (Kos, van den Brink, & Hagoort, 2012), and argument-structure (Kim et al., 2018). In fact, there seems to be a negative correlation between the respective magnitudes of the N400 and the P600 (Osterhout, McLaughlin, Kim, Greenwald, & Inoue, 2004; Tanner & Van Hell, 2014; see Chapters 2 and 3 for discussion). Further, the tradeoff between N400 and P600 effects seems to be consistent across experimental manipulations (Kos et al., 2012; Nickels, 2016; Tanner, 2019). For example, an individual with an N400-dominant profile may tend to maintain this dominant response to both lexical-semantic and agreement errors. In sum, the same stimuli can elicit different responses across individuals, but a given individual can elicit similar ERP responses across different types of linguistic anomalies. An investigation of what individual factors contribute to ERP response dominance may therefore inform us on the functional significance of the N400 and P600 components. For example, it has been suggested that reading span and working memory measures correlate with larger P600s (Bornkessel et al., 2004; Kim et al., 2018), congruent with the interpretation of the P600 as a P3b that indexes the updating of working memory for deviant stimuli (Sassenhagen et al., 2014).

However, without exception, all the studies that support the “tradeoff theory” (Kim et al., 2018) measure the N400 and the P600 using the same electrodes and most of them use adjacent time windows. It is therefore unclear whether the observed negative correlations are attributable to a genuine functional “tradeoff” or to an overlap between the N400 and the P600 components (see Brouwer & Crocker, 2017, for a discussion on how spatiotemporal component overlap is sometimes overlooked during ERP data interpretation).

While focusing on inter-individual variability in native (L1) speakers is relatively new, learners of a second language (L2) have traditionally shown more variability in their ERP profiles. Studies of sentence processing in L2 learners have identified main factors that account for these differences, such as age of acquisition, cross-linguistic similarity, and proficiency. In the next section we will address the relative importance of these factors, especially with respect to native-like attainment.

4. Sentence processing in the context of L2 learning

In L2 research, native-like proficiency is often associated with ELAN/LAN/P600 effects as opposed to N400 effects (Pakulak & Neville, 2011; Rossi et al., 2006). The relative importance of which factors promote native-likeness, however, remains under debate. We present a number of factors that have received support in the ERP literature.

4.1. L1ers vs. late L2ers: AoA and markers of nativelikeness

Age of acquisition (AoA) is a determining factor in reaching a high L2 proficiency level: the later one learns a second language, the less likely one is to achieve native-like proficiency. According to the critical period hypothesis (Johnson & Newport, 1989; Lenneberg, 1967; Newport, Bavelier, & Neville, 2001), there is a critical period in life when the brain can establish the neural connections that will enable optimal grammatical processing. Under the assumption that this critical period applies to all the languages one learns in one's lifespan, it follows that native-like processing in one's L2 is contingent upon early exposure. The critical period is generally limited to puberty (Birdsong, 2006; Vanhove, 2013), although a large-scale study has recently suggested that it extends until around 17 years (Hartshorne, Tenenbaum, & Pinker, 2018). Although AoA effects are discussed in more detail in Chapter 3, it is important to note that not all linguistic domains are equally affected by this factor: late L2 learners' lexical-semantic processing is generally quite comparable to L1 speakers' (Clahsen & Felser, 2006). In ERP research, late learners are generally found to elicit N400s not just in response to lexical-semantic incongruencies, but also to syntactic category violations, when native speakers would elicit and ELAN-P600 complex (Hahne & Friederici, 2001; Weber-Fox & Neville, 1996). In sum, late learners seem to rely on neurocognitive mechanisms that are reflected by the N400, regardless of the linguistic domain that is evaluated. However, the idea that late L2 learners and L1 speakers rely on different neuro-cognitive processes should be taken with caution as in the aforementioned studies AoA is confounded with language proficiency and experience. It is therefore difficult to attribute differences between L1 and L2 groups to AoA alone. In the next section, we will focus on studies that show that late L2 learners can converge toward L1 profiles with sufficient proficiency or thanks to cross-linguistic transfer between their L1 and their L2 (Caffarra, Molinaro, Davidson, & Carreiras, 2015; Steinhauer, 2014; Steinhauer et al., 2009).

4.2. L2 speakers converging on L1 speakers

Besides a decline in brain plasticity, age of acquisition effects can also reflect how entrenched the L1 system becomes through experience: the later one learns their L2, the stronger the impact of their L1 on their L2 grammar (Hernandez, Li, & MacWhinney, 2005). The influence of the L1 on an L2 suggests that differences and similarities between the two languages can lead to transfer effects (MacWhinney, 2005; White, 2003). Simply put, if similar syntactic structures are found in the L1 and the L2, then native-like processes may be observed in L2 learners for these specific structures. For example, Nichols and Joanisse (2019) observed that proficiency, not AoA, predicted the LAN amplitude in response to SC violations that are based on similar rules in French and English, while AoA was a better predictor of ERP responses to agreement errors (see Chapter 3 §1, for a critical discussion of their results).

When keeping cross-linguistic factors and AoA constant, proficiency itself seems to be a driving factor in native-like achievement. Proficiency evaluation is itself quite controversial (Leclercq, Edmonds, & Hilton, 2014), and can be assessed directly through offline measures (lexical decisions tasks, C-tests, or standardized tests such as the Test of English as a foreign language – TOEFL®), online measures (such as performance at an acceptability judgement task), or indirectly using the amount of exposure to an L2 (Ojima, Nakata, & Kakigi, 2005) – in this case, it can be conflated with AoA.

The convergence hypothesis (Steinhauer et al., 2009) suggests that distinct neurocognitive processing mechanisms (reflected by different ERP patterns) characterize proficiency levels, and that ultimately, highly proficient L2 learners may converge on native-like profiles. At early learning stages (low-proficiency level), L2 learners rely on frequency-based mechanisms to process syntax. As proficiency increases, L2 learners start applying grammatical rules (see also the Declarative/Procedural model, Ullman, 2001, 2004). In their review of morphological agreement and syntactic category processing, Steinhauer et al. (2009) observe that novice speakers tend to elicit N400s in response to grammatical errors, while intermediate speakers elicit a broadly distributed positivity, and advanced speakers a LAN/P600 complex similar to L1 speakers. These observations are compatible with Caffarra et al.'s meta-analysis (2015) which showed that proficiency was a reliable predictor for the presence of a P600 response to syntactic violations, while immersion predicted the presence of a LAN effect, and only violation type (SC violations) predicted the ELAN. Note that we can explain the ELAN as being an artefact in the majority of the L2 studies due to design flaws outlined above. However, Cafarra et al. only take the presence or absence of a significant difference at group level into account for their analyses. Although this measure is informative, we know that individuals vary immensely in the magnitude of their ERP responses to violations, and this may diminish effects at a group level.

4.3. Moving away from "ultimate L1 attainment": back to individuals

Even in native monolinguals, the magnitude of ERP responses may be modulated by proficiency (Pakulak & Neville, 2011; White, Genesee, White, King, & Steinhauer, 2006), which, in conjunction with L1 inter-individual variability described in the previous section, invites us to consider inter-individual variability in speakers, regardless of their linguistic status, instead of referring to a concept of "ultimate L1 attainment" (Hartshorne et al., 2018). Tanner, McLaughlin, Herschensohn, and Osterhout (2013) observed that proficiency predicted

L2 speakers' overall response magnitude, reflected by either a larger N400 or by a larger P600 (as the two components are negatively correlated). As we have seen above, response dominance profiles have also been observed in native speakers. One radical interpretation of the tradeoff hypothesis would be that individuals who display a specific ERP profile (who are, say, N400-dominant) would do so not only across experimental manipulations, but also across languages. However, interpreting L2 inter-individual variability data as definite evidence against the convergence hypothesis (Tanner et al., 2013) would be a misconception of what it really predicts. The convergence theory (Steinhauer et al., 2009, Steinhauer, 2014) does not predict that there is a specific LAN-P600 profile elicited by all native speakers, but that high proficiency drives responses that are similar to those of native speakers. One possibility, combining the tradeoff approach with the convergence theory, would be that proficiency is actually a by-product of motivation in a broader sense (Steinhauer, 2014). This concept includes motivation to learn a second language (Tanner, Inoue, & Osterhout, 2014), but also motivation to perform well during the experiment (Nickels, 2016). Individuals who are generally motivated to do well at a given task may exhibit distinct learning trajectories, which may lead to higher proficiency when AoA is kept constant. There is some evidence that ERP profiles in an L1 may predict success in learning an L2. Qi and collaborators (2017) manipulated semantic incongruencies in their participants' first language and observed that N400-dominant responses predicted successful initial vocabulary learning of an artificial language, while P600-dominant responses predicted successful initial learning of grammatical structures.

In sum, understanding what causes the variability observed in both L1 and L2 speakers may be crucial in further characterizing the functional significance of ERP components elicited during sentence processing.

5. The present studies

5.1. Objectives and research questions

The primary objective of this thesis is to re-evaluate the time-course of sentence processing, and in particular, to clarify the timing and nature of the mechanisms underlying SC identification. To this end, the present studies introduce a novel paradigm in French that enables the investigation of two specific questions:

(i) Are SCs analyzed first?

(ii) Is lexical-semantic processing contingent upon phrase-structure building?

Second, the examination of individual ERP responses to SC violations will address two additional issues related to inter-individual variability that are currently under debate, but have mostly been discussed in the context of morpho-syntactic agreement:

(iii) Is the LAN a by-product of a grand-averaging of ERP responses?

(iv) Do individuals exhibit either N400 or P600-dominant profiles in response to syntactic category violations?

A third objective is to investigate the underlying mechanisms of SC and lexical-semantic processing in late L2 learners of French (L1: English). Group comparisons, as well as the evaluation of the relative importance of several predictors related to AoA, proficiency, and exposure, will address the following question:

(v) Do late L2 learners converge toward L1 profile with sufficient proficiency, as predicted by the convergence hypothesis (Steinhauer, 2014)?

Finally, by focusing on Quebec French, this project will contribute to bringing cross-linguistic evidence to the neurolinguistics literature.

5.2. Study 1

This study uses an original paradigm that systematically manipulates contextual semantic priming and target-word syntactic category. Syntactic category violations were introduced by replacing a noun with a verb (and vice versa), while ensuring that the context words immediately preceding the target noun or verb were matched, so that any ERP baseline issues that undermined previous studies were avoided. Four main conditions involving both noun and verb targets were created by manipulating these syntactic categories (correct/syntactic category violations), and lexical-semantic anomalies (primed/unprimed).

Contra Friederici (2002), we hypothesized that: (i) there would be no difference between syntactically correct targets and syntactic category violations in the early (i.e. in the 100-300 ms) time-window, but that syntactic category violations would elicit a LAN instead, and (ii) semantically anomalous sentences would elicit N400 effects *even in the presence* of a syntactic category violation. Exploration of individual datasets will test whether component overlap between and within participants may explain: (iii) an "illusion" of a LAN (Tanner, 2015), and (iv) a tradeoff between the P600s and preceding negativities.

5.3. Study 2

The second study examines these issues further, and develops some questions related to inter-individual variability by considering L2 learners in addition to L1 speakers of French. The introduction of Random forests, a statistical approach that is novel to ERP research, will allow us to test the convergence hypothesis by ranking the relative importance of several continuous predicting variables. We hypothesise that (v) proficiency will modulate the ERP profiles of L2 learners to a larger extent than AoA and may account for some variability *even within native speakers*.

Chapter 1. Methodological considerations

The following Chapter describes information about participants, experimental design, stimuli evaluation, and procedures that are relevant to both chapters that follow it. In particular, we expand on participants' demographics and the behavioral measures that we used to evaluate their proficiency in French. A second area of interest concerns the evaluation of our stimulus materials in initial rating studies, in order to ensure that the experimental design would be appropriate to investigate the online processing of syntactic categories without undesirable confounds.

1. Participants

This project required the recruitment of 225 participants in total (no participant contributed to more than one experiment – behavioral or EEG). Eighty monolingual speakers of Quebec French were recruited to complete stimuli evaluation tasks, namely acceptability ratings (n = 40) and cloze task (n = 40). Those tasks were completed at the *Laboratoire de recherche sur l'acquisition et le traitement du langage* of the *Université de Montréal* (UdeM) (Dr. P. Royle, dir.), and all participants provided informed consent by signing forms approved by the *Comité d'éthique de la recherche en Santé* at the UdeM. These participants were students who volunteered in exchange for course credit. In addition, relational priming was evaluated through an online survey (n = 60) that did not require any personal information, so this task was exempt from evaluation by an ethics review board.

Forty monolingual speakers of Quebec French (L1 study) and 45 English speakers who learned French after puberty (L2 study) participated in the EEG main experiment that was conducted in the visual modality. All participants were tested at McGill University's

Neurocognition of Language Lab (Dr. K. Steinhauer, dir.); consent forms, materials, and procedures were approved by the Ethics Review Board of the Faculty of Medicine, McGill University, and the *Comité d'éthique de la recherche en santé,* UdeM. Individuals who volunteered for the EEG recording session were compensated $50 for their time. Participants were right-handed, as determined by the abbreviated version of the *Oldfield Handedness Inventory* (Oldfield, 1971, adapted to French for native monolingual speakers), and reported normal or corrected-to-normal vision and no reading or neurological disorders (e.g., dyslexia). Particularly relevant to our study on sentence processing in L2 learners was information gathered about their language background, use, and proficiency. The behavioral measures that we employed are described below, along with descriptive statistics for each group of participants in the ERP studies (L1 and L2).

1.1. Language-background measures

All participants who took part in EEG/ERP experiments filled out a questionnaire on Google Forms[8] containing questions about the participants' demographic information: age, education, and reading habits. More specific language background measures were collected by means of a pen-and-paper questionnaire developed at the UdeM Lab. The detailed questionnaire is available in Appendix 1. Three variables of interest were quantified from the participants' responses to the questionnaire. Age of Acquisition (AoA) of French is directly relevant to the critical period hypothesis, but late L2-learners often relate it to their first "classroom" experience, which we wished to distinguish from the onset of regular exposure to the target language. A second variable called Age of Exposure (AoE) corresponds to the age at

[8] www.docs.google.com/forms/

which participants started to communicate in French on a daily basis. For both variables, a default value of zero was attributed to native speakers. Third, daily exposure reflects how much time (in percentage) the participants use French every day in the following situations: at home, with friends or during social activities, and at school or work.

1.2. Language proficiency and working memory measures

Two independent measures of proficiency in French were administered using Paradigm Player (Perception Research Systems, 2007): LexTALE_FR and a cloze test. LexTALE stands for *Lexical Test for Advanced Learners of English* and is a lexical decision test that was originally designed to assess participants with an advanced level of English in experimental settings. It is a good predictor of vocabulary knowledge and correlates with proficiency measures such as the TOEIC score (Lemhöfer & Broersma, 2012). LexTALE_FR is the French version, normed with over 500 native speakers and 200 second language learners (Brysbaert, 2013). It comprises 60 trials and takes 3.5 min to complete on average. Instructions, materials, and the response key to this test (retrieved from: http://crr.ugent.be/archives/921) are in Appendix 2.

The cloze test was developed by Tremblay and Garrison (2010) to assess the proficiency of “intermediate” to “advanced” adult L2 learners of French for psycholinguistic research purposes. The base text is a non-academic article from the French newspaper *Le Monde* containing 314 words, 45 of which are deleted (23 are content words). Participants are asked to fill in the blanks by choosing between four different options for each word. There is no time limit for this test, however participants cannot correct their choice once it has been made. The scores are calculated as percentage correct. What makes this test particularly interesting for the purpose of this research project is that it was piloted with both native speakers from France and Québec. The original study by Tremblay and Garrison (2010) demonstrates that the scores increase with proficiency, but do not quite reach ceiling (the most proficient learners had an average score of 82.9%; range: 73.3–93.3%).

1.3. Speakers of French in Montreal

Montreal is located in the province of Quebec where French is the official language (Bill 101 – Charter of the French language, retrieved from: http://www.legisquebec.gouv.qc.ca/fr/showdoc/cs/C-11). However, due to historical, geopolitical, and cultural reasons, English is very present: Montreal thus hosts a bilingual and even multilingual environment (data retrieved from: http://ville.montreal.qc.ca/portal/page?_pageid=6897,67887637&_dad=portal&_schema=PORTAL). Therefore, even French native speakers who grew up off the Island of Montreal in a monolingual family (and who constitute most of our L1 group) end up being exposed to some English, and therefore understand it much better than individuals who grow up in, say, France. This being said, francophone and anglophone communities tend not to blend: there can be a lack of communication, sometimes referred to as the "two solitudes" (MacLennan, 1945). It follows that, with some exceptions, members of the anglophone community who learn French (even at high proficiency) are not that much exposed to it in their daily lives. This is reflected by the relatively low amount of daily exposure to French in our L2 group, which may also be a limitation in our study, since exposure may, among other linguistic aspects, impact lexical development (Barriere, 2010). Table I summarizes participants' demographics and language measures split by group. Native and L2 participants were selected so the two groups would be as similar as possible with respect to their age, level of education, reading habits (which is an indirect measure of their reading skills), handedness, and working memory (measured with a forward and backward digit span): there were no difference between groups on these measures. Differences in the LexTALE_FR score and the cloze test reflect differences in proficiency, and the wide range observed at both tests will allow us to consider these factors as continuous variables in the ERP analyses. Daily exposure does not overlap between L1 and L2 speakers, which may create a confound with Group.

Table I. Demographics and language background for each group (L1-L2). Signification differences in bold.

Measure	L1 French (*n* = 36, 19F)		L2 French (*n* = 41, 26F)		*t*(df)	*p-value*
	M (*SD*)	Min – Max	*M* (*SD*)	Min – Max		
Age (years)	27.00 (5.47)	22 – 40	25.76 (4.08)	25 – 34	1.11 (76)	0.270
Education (years)	15.35 (1.92)	13 – 18	14.40 (2.86)	10 – 18	1.77 (76)	0.081
AoA (years)	**0.00 (0.00)**	–	**12.44 (1.94)**	**10 – 18**	**43.09 (40)**	**< 0.001**
AoE (years)	**0.00 (0.00)**	–	**17.96 (3.50)**	**12 – 27**	**34.39 (40)**	**< 0.001**
Daily exposure[a]	**88.59 (10.69)**	**60 – 100**	**17.67 (13.90)**	**3 – 60**	**25.29 (76)**	**< 0.001**
Reading Habits[b]	3.88 (0.81)	1 – 4	3.98 (0.83)	3 – 5	-0.51 (76)	0.609
Handedness	81.29 (16.01)	50 – 100	82.44 (16.88)	40 – 100	-0.31 (76)	0.755
LexTALE_FR[c]	**89.43 (4.04)**	**78 – 95**	**58.80 (11.03)**	**16 – 83**	**16.97 (57)**	**< 0.001**
Cloze test[c]	**68.33 (12.47)**	**33 – 91**	**39.65 (17.56)**	**11 – 89**	**8.09 (69)**	**< 0.001**
Reliable digit span[d]	9.94 (2.41)	3 – 13	10.00 (2.65)	5 – 16	-0.09 (70)	0.926

[a] Percentage of daily exposure to French estimated since age 18
[b] On a scale from *0 – never reads* to *5 – reads a lot*
[c] Scores computed in percentages
[d] Task adapted from Soylu (2010). The sum of the longest string of digits recalled correctly twice, under both forward and backward conditions (Greiffenstein, Baker, & Gola, 1994).

2. Presentation of the experimental design

In order to address the methodological shortcomings of previous studies identified in the Introduction and provide an accurate account of syntactic and lexical-semantic integration during sentence comprehension, the experimental design systematically manipulates syntactic categories and lexical-semantic priming, while remaining perfectly balanced (Table II). The starting point in the development of this design is the homophony and homography, in French, of clitic pronouns (e.g., *le* /lœ/ 'him') and definite determiners (e.g., *le* 'the'): both *le* + verb and *le* + noun are grammatical combinations. Taking advantage of this feature, pairs of correct target sentences (marked ✓✓ in Table II) were created with the following characteristics:

a. Context sentences introduced (i) antecedents (i.e. referents) to the clitic pronouns and definite determiners, and (ii) a prime for target words (e.g., *hockey* primes *plaquer*);

b. Matched pairs of verbs consisted in (i) control verbs that require a verb phrase as a complement, so we could introduce a clitic pronoun and a verb target (e.g., *Elles osent le* + verb) or (ii) transitive verbs that require a noun phrase as a complement (e.g., *Elles ôtent le* + noun);

c. The sentences also contained matched pairs of verb-noun target words (*plaquer* and *crapaud*). Note that in contrast to Isel et al (2007), our use of the infinitive introduces affixal markers for verbs.

Table II. Sample stimuli for the four experimental conditions.

Syn	Sem	Context sentence	Experimental sentence			
			Subject + Verb	Clitic / Determiner	Target	Prepositional phrase
✓	✓	*Marie et Jeanne jouent au hockey avec leur copain.*	*Elles osent*	*le*	*plaquer*	*sur le côté.*
		Mary and Jane are playing hockey with their friend$_{\text{MASC}}$.	They$_{\text{FEM}}$ dare	him	tackle	on the side.
		Marie et Jeanne vont au marais avec leur copain.	*Elles ôtent*	*le*	*crapaud*	*sur le côté.*
		Mary and Jane are going to the swamp with their friend$_{\text{MASC}}$.	They$_{\text{FEM}}$ remove	the$_{\text{MASC}}$	toad$_{\text{MASC}}$	on the side.
✗	✓	*Marie et Jeanne vont au marais avec leur copain.*	*Elles osent*	*le*	**crapaud*	*sur le côté.*
		Marie et Jeanne jouent au hockey avec leur copain.	*Elles ôtent*	*le*	**plaquer*	*sur le côté.*
✓	✗	*Marie et Jeanne vont au marais avec leur copain.*	*Elles osent*	*le*	*?plaquer*	*sur le côté.*
		Marie et Jeanne jouent au hockey avec leur copain.	*Elles ôtent*	*le*	*?crapaud*	*sur le côté.*
✗	✗	*Marie et Jeanne jouent au hockey avec leur copain.*	*Elles osent*	*le*	**crapaud*	*sur le côté.*
		Marie et Jeanne vont au marais avec leur copain.	*Elles ôtent*	*le*	**plaquer*	*sur le côté.*

In the setup exemplified in Table II, outright syntactic category violations were created by simply swapping the noun and verb complements. Further, swapping the introductory sentences removed the semantic prime and made the sentences semantically anomalous. The design ensured that we had a balanced design: both targets and contexts immediately preceding targets appear in each condition. In total, four conditions were created by manipulating these two dimensions: syntactic category (Syntax: grammatical/ungrammatical), and semantic priming (Semantics: correct/anomalous), and by combining them using a Latin-square design.

This experimental design is among the very few that introduce outright syntactic category violations and lexical-semantic anomalies (but see Bowden et al., 2013; Steinhauer et al., 2009 for examples in Spanish and English) as opposed to the seminal ELAN (Friederici et al., 1993) and "semantic blocking" (Friederici et al., 1999) studies. In order to maintain a balanced design, it was crucial to minimize differences between pairs within each condition, so that observed ERP effects could be really attributable to syntactic and semantic anomalies rather than differences that may arise from lexical characteristics, differences in acceptability, or semantic-pragmatic context effects linked to specific stimulus items. In the following, procedures to control for these factors are described.

3. Creation and evaluation of stimuli sentences

3.1. Matching sentence pairs

Since our design relies on pairs of sentences in each condition, any difference within each pair could impact the ERP responses. For example, a more frequent target word in one condition would decrease the N400 component in response to this word, and this effect could interact with negativities predicted in response to lexical-semantic anomalies and syntactic category violations (Van Petten & Kutas, 1990). Special attention was also paid to the context preceding the target, following S&D2012's argument that any ERP differences elicited by the context words (e.g., an N400 effect) can either spill over or induce artefacts in the time-window of interest. Orthographic differences (Carreiras, Perea, Vergara, & Pollatsek, 2009), differences in length (in French: Babin, 1998; Labelle, 2001), and as complexity variability in the syntactic structure of sentences (e.g**.**, Featherston, Gross, Münte, & Clahsen, 2000) could also create unwanted ERP differences either before or during target word presentation. For these reasons, we controlled for all these dimensions when creating our sentence pairs.

We first selected 20 pairs of control and transitive verbs (such as *oser* and *ôter*, respectively). Homophony between determiners and clitic pronouns already made the context immediately preceding the target constant, but in order to eliminate any possible lexical effects within the contexts, we ensured the verbs perfectly matched in syllable length (M = 1.6, SD = 0.88), and did not differ significantly in number of phonemes, syllables, characters, and log-transformed frequency (Table III).

Table III. Paired t-test results comparing control verbs and transitive verbs.

Lexical dimension	*Mean* (*SD*)		Paired t-tests results	
	Control verbs (e.g., *oser*)	Transitive verbs (e.g., *ôter*)	t(df)	*p* value
Num. Phonemes [a]	3.8 (1.23)	3.8 (1.32)	< .001	1
Num. Characters	6.45 (1.36)	6.20 (1.51)	1.097	.287
Frequency [b]	2.35 (0.94)	2.80 (1.24)	-1.360	.190

[a] Number of phonemes are based on Québec French phonology.
[b] All lexical measures were retrieved from Lexique.org. Frequencies are log-transformed values from lemma frequencies for French film subtitles.

Selecting control verbs that only take a verb as a complement was particularly challenging, as nouns can be "coerced" after most control verbs as complements or adjuncts. We first made a list of 28 control verb candidates and had ten participants suggest an appropriate continuation for the following sequence: NP + control verb + *le*/*la*/*les* (…). The 20 verbs that were ultimately selected were followed by a verb at least 80% of the time. Following the recommendations of the thesis committee, we ran another *post hoc* test where we specifically asked 16 participants to place a *noun* after each NP + control verb + *le*/*la*/*les* sequence whenever possible. Six additional participants rated the acceptability of the sentences that were completed with a noun (1– perfectly acceptable; 4– totally unacceptable). The results showed that a noun complement could be coerced upon five main verbs over 50% of the time, and that the resulting sentences were judged to be acceptable ($M < 2.01/5$). In addition, two verbs (*partir* 'to leave' and *venir* 'to come') could be followed by adjunct nouns (e.g., *Marie part le soir* 'Marie leaves at night'). Due to the repetitive nature of the experimental design, we did not expect our participants to expect adjuncts following any of the main verbs, but we acknowledge that this possibility may still affect the ERPs.

Each of the twenty control-transitive verb pairs selected for eight verb and noun targets, respectively, to create a grand total of 320 sentences (160 pairs). Verbs and nouns that belonged to the same pair matched in syllable length, number of phonemes, and log-transformed frequency. Note that since half the target nouns were plural and carried an additional *-s* plural marker (this *–s* is silent in French), they were typically longer than verbs on number of characters (Table IV). Matching for number of phonemes over number of characters was prioritized in order to keep stimulus duration as equal as possible in the auditory modality.

Table IV. Paired t-test results comparing target verbs and nouns.

Lexical dimension	*Mean* (*SD*)		Paired t-tests results	
	Target verbs (e.g., *plaquer*)	Target nouns (e.g., *crapaud*)	*t*-value	*p*-value
Syllable length	2.10 (0.04)	2.09 (0.04)	- 0.706	.438
Num. phonemes	5.12 (0.08)	5.06 (0.08)	- 1.164	.246
Num. characters	6.52 (0.09)	6.94 (0.08)	4.784	< .001
Frequency	1.31 (0.77)	1.21 (0.74)	- 2.122	.103

To license the use of definite determiners and clitic pronouns, and to enhance sentence naturalness, experimental sentences were preceded by introductory sentences: the referent for the clitic or determiner was also a prime for the target word. To avoid any effect of length on working memory (e.g., Piai, Meyer, Schreuder, & Bastiaansen, 2013), we matched the contexts in number of words (± 1 word) and syntactic structure whenever possible without compromising the semantic integrity of the sentences (see complete list in Appendix 3). Referential ambiguity was avoided by counter-balancing the gender and number of subject pronouns (e.g., in Table I, *Marie et Jeanne … Elles* 'they$_{\text{FEM.PLUR}}$ denote a feminine plural referent in French) vs. clitic pronouns (e.g., *le* 'the' is masculine singular). Finally, prepositional phrases concluded the sentences to avoid possible ERP wrap-up effects on target words (Hagoort, 2003b; Just & Carpenter, 1992; Stowe, Kaan, Sabourin, & Taylor, 2018).

After creating syntactically and semantically correct sentence pairs, we swapped the target words and introductory sentences to create syntactic category and lexical-semantic anomalies. In the following, we demonstrate how we ensured both that correct conditions would be equally acceptable, and incorrect conditions equally unacceptable, to native French speakers.

3.2. Acceptability judgements

To maintain a balanced design, the sentences within each pair should not differ in their acceptability. We focused on “Correct” sentence pairs and pairs that “Combined” syntactic category and lexical-semantic anomalies. All 160 sentence pairs with verb types (“Transitive” vs. “Control”) in the Correct and Combined conditions were separated into four experimental lists, using a Latin square design, so that every participant would only see one sentence per pair and condition. Forty undergraduate students of the *UdeM* participated after giving informed consent. They were asked to rate their acceptability on a scale from 1 — *acceptable* to 4 — *unacceptable*. Repeated measure ANOVAs revealed a main effect of Condition (2 levels, Correct: $M = 1.69$ ($SD = .44$) and Combined: $M = 3.64$ ($SD = .30$); $F(1, 159) = 3782.08$, $p < .001$. There was no significant effect of Verb type (2 levels, Control: $M = 2.67$ ($SD = 1.06$) and Transitive: $M = 2.66$ ($SD = 1.04$), $F(1, 159) = 0.427$, $p = .514$) and no interaction of these two factors ($F(1, 159) = 1.689$, $p = .196$). These results suggest that (a) participants judged the sentences appropriately, and (b) did not judge the sentences within pairs differently, suggesting that our design is balanced with regards to sentence acceptability.

3.3. Context effects

Because the context sentences introduce a semantic-pragmatic context prior to the target words, we checked whether this could influence target predictability within sentence pairs. Differences in the degree of priming between prime and target within a sentence pair could also affect the N400: more priming could, for example, lead to a reduced N400 in the semantically correct conditions (Kiefer, 2002). To evaluate priming magnitudes in semantically correct sentences, pairs of primes and targets (*e.g.,* "hockey" and "tackle") were presented to sixty different French speakers who were asked to rate their relatedness on a scale from 1 — *not related* to 5 — *very related*. Paired t-tests revealed no difference between our prime-target pairs across target types (Verbs: $M = 3.25$, $SD = 1.04$; Nouns: $M = 3.39$, $SD = 1.10$; $t(159) = 1.353$, $p = .18$).

Second, since N400 effects can also be observed in response to unexpected words given a sentence or discourse context (Kutas & Hillyard, 1980), we evaluated semantic-pragmatic effects on the predictability of the targets by running a cloze test. We removed the target word from all 160 sentences pairs in the correct and semantically anomalous conditions, and divided them into four lists using a Latin square design. Forty French-speaking participants were asked to complete the sentences with what they thought was the most appropriate word. As expected, the cloze probabilities for semantic anomalies were almost always zero, so we only used the responses on the correct sentences for our evaluation. We then calculated the probability of the target words to be used based on sentence-completion results on a range from 0 – the target was never suggested as best completion – to 1 – the target was always produced as completion. Paired t-test revealed that target words selected by a control verb ($M = .10$, $SD = .18$) were less probable than targets following a transitive verb ($M = .17$ ($SD = .26$) ; $t(159) = 2.774$, $p = .006$). It is also worth noting that although the association between prime and target word was rated as being quite high (above 3/5), cloze probabilities were very low for these sentences, potentially leading to a moderate reduction of the N400 effect in the primed sentences. Since we could not match items on this specific dimension without impeding the other lexical, formal, acceptability, and relatedness dimensions that we had carefully controlled for, we opted to keep these stimuli and include cloze probability as a random slope in our statistical models (see Chapter 2 §2.5).

It is also worth noting that some of the transitive verbs are more constraining than others with respect to the semantic category of the noun they select. For example, *boire* 'to drink' specifically selects a liquid: there is a local priming effect in addition to the priming introduced in the sentence context, that may further affect the differential ERPs elicited by ungrammatical sentences vs. grammatical ones. This priming is largely reflected by the cloze probability measures that were collected and included as a random slope in the mixed-effect models. Further inspection of ERP effects divided across levels of semantic constraint suggests that it is unlikely that this dimension affected our syntactic and semantic manipulations differentially (cf. Appendix 4).

4. EEG testing

4.1. Experimental procedure

Participants were tested in a 2.5-hour session in the Neurocognition of Language Lab at McGill University. First, they completed the language background, sleep, reading habits, and demographics questionnaires (about 15 min total). Then, the EEG session started: after cap placement (about 30 min), participants were seated in a chair 80 cm away from a computer monitor and read the sentences while their EEG was recorded. The sentences were presented in rapid-serial-visual presentation mode (each word: 300 ms display + 200 ms blank screen), and words appeared in white 30-point Arial font on a black background. At the end of every trial, a “???” prompt appeared and remained on screen until participants scored sentence acceptability on a scale of 1 to 5 (1: totally acceptable – 5: totally unacceptable) by pressing a button on a response box. After participants responded, a second prompt “!!!” would appear for a 1800 ms interval during which participants were encouraged to blink their eyes (in order to avoid eye blinks during sentence presentation). The experimental sessions started with two short practice blocks of six sentences each, while the 400 experimental trials were divided into eight blocks of 50 trials each, separated by short breaks. After the EEG experiment, participants completed the digit span task, the cloze test, and the lexical decision test in French.

4.2. EEG recording and preprocessing

EEG was recorded continuously from 25 Ag-Cl active shielded electrodes mounted on an EEG cap (Waveguard™ original, *ANT Neuro*, Netherlands) according to the 10-20 system (Jasper, 1958) at the following sites: Fp1, Fpz, Fp2, F7, F3, Fz, F4, F8, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1, Oz, O2. All EEG electrodes were referenced online against the right mastoid. An electrode placed halfway between Fpz and Fz served as ground. Impedances were kept below 5 kΩ. EEG was recorded at a sampling rate of 512 Hz with a 0.001 — 100 Hz online filter.

Data were analyzed using *EEGLab* (Delorme & Makeig, 2004) and *ERPLab* (Lopez-Calderon & Luck, 2014). Continuous data were re-referenced offline to average mastoids, and bandpass-filtered with .1 and 40 Hz cut-off frequencies (IIR Butterworth filter). After epoching the data from -1000 to 2000 ms relative to the onset of the target word, we rejected data that exceeded a peak-to-peak threshold of 75 μV (in 100 ms steps). We then visually inspected the remaining epochs and deleted ones that still appeared to be affected by artefacts. The analyses were then performed on shorter epochs of -200–1800ms with respect to target word onset, and baseline corrected (-200–0ms). Event-related potentials were quantified using the average activity over representative time-windows (details in Chapters 2 and 3) for every single observation (Chapter 2) or on aggregated data over sub-conditions (Chapter 3).

5. Statistical analyses of individual differences

5.1. On selecting an appropriate technique

There are many ways to investigate individual differences in the ERP responses of L1 and L2 participants. One possible approach is to first observe individual profiles and test if they correlate with certain predictors, such as AoA, proficiency, motivation, or working memory (Kim et al., 2018; Tanner et al., 2013). This approach has led to what we described in the Introduction as the "tradeoff hypothesis": a response dominance (toward N400 or P600) is calculated, and regressions (Tanner, 2019) are used to account for it. Of course, now that it has been established that inter-individual variability has to be taken into account when developing models of sentence processing, we should be able to determine which cognitive, linguistic, environmental factors best account for individual responses. In other words, we need to find a way to cluster individuals back into groups. This objective has been particularly challenging, partly because many of the aforementioned factors are correlated. One way to circumvent this issue is to apply dimension reduction to the data: for example, by performing a principal component analysis to the variables of interest. Tanner (2019) extracted two principal components that correlated with working memory and language proficiency measures, respectively, and showed that none of them had any effect on the ERPs that were elicited by agreement violation errors. One issue with principal components is that one can simply *assume* that because they correlate with specific factors (e.g., lexical decision task and C-test results), they represent the latent variable that underlies them (e.g., proficiency). However, it is in fact very unclear what the extracted PCs actually represent: in our example, it could be general proficiency or vocabulary knowledge. In the present thesis, we propose an alternative way of analysing the data that allows us to (1) deal with multicollinearity with *Random forests*, and (2) cluster the data in groups that can be predicted by the most important factors with inspection of decision trees.

5.2. Random forests

The random forests methodology was adapted from Tomaschek, Hendrix, and Baayen (2018), using unconditional variable importance with the *ranger* package in R (Wright & Ziegler, 2017). This method is known to be well suited for dealing with correlated variables (Strobl, Malley, & Tutz, 2009) and does not inflate the importance of continuous variables (e.g., proficiency) over categorical ones (e.g., group). To estimate variable importance, the algorithm randomly selected subsets of the data and modeled the effect of each predicting variable in every subset. Accuracy of each prediction was compared to the remaining observations. Strength of a predicting variable was calculated by randomly permuting its levels and thus erasing its importance: a predictor is deemed important if the model becomes worse (Strobl, Boulesteix, Kneib, Augustin, & Zeileis, 2008).

Then, we used conditional inference trees to illustrate how the most important variables interact (Hothorn, Hornik, & Zeileis, 2006). Trees predict the value of continuous variables (ERP amplitudes) from a set of continuous or categorical predictor variables, using recursive binary partitioning. They provide estimated split points at which the nodes separate between two groups with different outcomes. The splitting criteria are calculated using the permutation-test framework (Hothorn et al., 2006). For each possible split, the test-statistic value is calculated under a certain label rearrangement: if they are interchangeable, the splitting value is not relevant.

6. Conclusion

This chapter introduced the general methodology relevant to both ERP studies that will be presented. Stimuli description was particularly detailed because we will argue in the next chapter that their meticulous development was key to obtain clear and unambiguous ERP data to answer our research questions. In Chapter 2, we make use of this balanced design to investigate the time-course of SC identification and lexical-semantic processing in native speakers of French. In particular, the data will allow for a critical reevaluation of the "syntax-first" approach to sentence comprehension.

Chapter 2. ERP evidence against "syntax-first" approaches to sentence processing

Lauren A. Fromont [a,b,*]***, Karsten Steinhauer*** [b,c]***, Phaedra Royle*** [a,b]

[a] École d'orthophonie et d'audiologie Université de Montréal, Montréal, Canada

[b] Centre for Research on Brain, Language and Music, Montreal, Canada

[c] School of Communication Sciences and Disorders, McGill University, Montreal, Canada

This Manuscript has been submitted to PLoS ONE and is currently under revision.

Abstract

In this event-related potential (ERP) study we reevaluate syntax-first approaches to sentence processing by implementing a novel paradigm in French that included correct sentences, pure syntactic category violations, lexical-semantic anomalies, and combined anomalies. This balanced design systematically controlled for target word (noun vs. verb) and the context immediately preceding it. Group results from 36 native speakers of Quebec French revealed that, up to 300ms, the ERPs elicited by syntactic category violations were perfectly aligned with the ERP responses to correct sentences, showing that there is no early activation reflecting syntactic category identification. Instead, in response to all anomalous conditions, we observed an N400 followed by a P600. While combined anomalies yielded additive effects of syntactic category and lexical-semantic anomalies on the N400, and a large P600 effect similar to the one observed in the pure syntactic condition. These results provide strong evidence against the hypothesis that (i) syntactic categories are processed first, and (ii) that syntactic category errors "block" lexical-semantic processing. Further, the N400 effect in response to pure syntactic category violations reflects a mismatch detection between a predicted word-stem and the actual target. This mechanism takes place simultaneously (and potentially in parallel) with lexical-semantic processing. The interactive effect on the P600 reveals that the same neurocognitive resources are recruited for syntactic and semantic integration, both promoted by the implementation, in our design, of an acceptability judgement task. Additional analyses of individual data complemented these observations: during sentence processing, participants did not rely on one single cognitive mechanism reflected by either the N400 or the P600 effect but on both, suggesting that the biphasic N400-

P600 ERP wave *can* indeed be considered to be an index of phrase-structure violations in most individuals, at least if they are realized on content words.

1. Introduction

Sentence comprehension, as effortless as it seems, is contingent upon a rapid analysis of each incoming word along several linguistic dimensions. Within a few hundred milliseconds, we identify the syntactic category of each word in order to build larger phrases (e.g., a noun is combined with its determiner to build a determiner phrase) and use its lexical-semantic properties to integrate it with the sentence and discourse context. Event-related potentials (ERPs), thanks to their excellent temporal resolution, allow us to study the time course of semantic and syntactic processing during on-line sentence comprehension. However, despite decades of research, the relative timing of cognitive processes underlying these two linguistic dimensions is still controversial. Friederici's serial and modular account of sentence processing adopts a "syntax-first" approach (Friederici, 2002; Friederici & Singer, 2015) where syntactic category identification of each incoming word occurs first and conditions further lexical-semantic and morphological analyses on a that word. While this framework has largely dominated the field since the mid 1990s, it has faced contradictory data (Dikker, Rabagliati, & Pylkkänen, 2009; Lau et al., 2006; Van den Brink & Hagoort, 2004) and a critical review of most previous studies supporting the "syntax-first" approach (Steinhauer & Drury, 2012). In that review, Steinhauer and Drury (2012, hereafter S&D) outline major methodological flaws observed in the majority of ERP studies investigating syntactic category and lexical-semantic processing, and call for cautious interpretation of ERP results related to them. Following S&D's arguments, we investigate the hypothesis that experimental designs are potentially responsible for the contradictory evidence for syntax-first models and propose a novel and improved experimental design to reevaluate the time-course of syntactic-category identification and lexical-semantic processing. By doing so, we will contribute to informing and extending language-processing models.

Friederici's "syntax-first" approach (Friederici, 1995, 2002, 2011) is arguably the most influential ERP-based neurocognitive model of language processing to date. It posits three phases, defined here along with the ERP components that index them:

Between 100 and 300 ms after stimulus onset, automatic identification of syntactic categories (noun, verb, etc.) takes place to generate initial syntactic representations. This phase is labelled *local phrase-structure building*.

The early left anterior negativity (ELAN), appears around 150 ms after stimulus presentation in response to syntactic category violations (Friederici, 1995).

Between 300 and 500 ms after stimulus onset, lexical-semantic information and morpho-syntactic relations are processed in parallel.

The N400 component is a negative deflection peaking around 400 ms after stimulus onset, usually observed at centro-parietal sites (e.g. Kutas & Federmeier, 2011; Kutas & Hillyard, 1980). The N400 has generally been associated with lexical-semantic processing: more specifically, and important for our design, it is sensitive to both priming effects (priming leads to a reduction of the N400), and cloze probability (lower cloze probability increases the N400);

The left anterior negativity (LAN) is a negative wave that peaks around 400 ms, often with a left-lateralized and frontal distribution (in the visual modality). It has been observed in response to morpho-syntactic errors, such as agreement violations (Kutas & Hillyard, 1983; Molinaro et al., 2011; Royle, Drury, Bourguignon, & Steinhauer, 2012).

Between 500 and 1000 ms after stimulus onset, information from various processing streams (e.g., syntactic and semantic information) are integrated, and, if necessary, a revision of the first analysis is initiated.

The P600 component is a large, long-lasting posterior positivity, starting around 500 ms, that is elicited by syntactic violations (Osterhout & Holcomb, 1992), as well as syntactically complex or temporarily-ambiguous sentences (Kaan, Harris, Gibson, & Holcomb, 2000). Some suggest that the early P600, observed between 500 and 750 ms – and distributed over midline electrodes (Molinaro et al., 2011) – reflects the reactivation of contextual information in order to process syntactic integration. The late P600, observed at parieto-occipital sites between 750 and 1000 ms or beyond, is assumed to reflect general sentence reanalysis and repair (Molinaro, Vespignani, et al., 2011) as well as controlled processes related to decision-making and categorization, as it is highly modulated by the presence or absence of a task (Royle et al., 2013) and by experimental design (Hahne & Friederici, 1999).

This model predicts that in Phase I, syntactic category information enjoys a special status and is used prior to lexical-semantic information for sentence comprehension. The ELAN effect that is specifically observed in response to syntactic category errors is the cornerstone of Friederici's proposal (Friederici, 2002). This effect is time-locked to the syntactic category cue at word onset or offset, for example inflectional morphology (Friederici, Gunter, Hahne, & Mauth, 2004b; Friederici, Hahne, & Mecklinger, 1996; Hagoort, Wassenaar, & Brown, 2003), and reflects automatic linguistic processes that are putatively independent from attention (Hahne & Friederici, 1999) and task demands (Hahne & Friederici, 2002). However, ELAN effects are not robust: for example, less than half of the studies reviewed by S&D report one. These authors report major methodological issues associated with ELAN studies. In particular, ELANs have been mostly evidenced using one specific context manipulation paradigm (e.g. *Die Bluse wurde oft gebügelt*, 'The blouse was often ironed' vs. *Die Bluse wurde am *gebügelt*, 'The blouse was on-the *ironed' Friederici, 1995; Friederici et al., 1993; Friederici et al., 1999) which became a standard paradigm for ELAN studies. Insertion of this preposition-article contraction *am* ('on-the') creates a syntactic category violation on the underlined target past-participle verb (purportedly indexed by the participle verb prefix *ge*-) because a noun would be expected rather than a verb, resulting in an ELAN effect on the target word. According to S&D, context-manipulation paradigms can lead to artefactual context effects in the time-window of the target word. For example, an N400 difference in the baseline window is observed when comparing a context ending with a content word in the correct condition (*oft* 'often'), to a context ending with a function word in the syntactic category violation condition (*am* 'on-the'). Baseline correction would then shift the whole ERP waveform in the correct condition towards a positivity, and in the incorrect one towards a negativity, leading to an early, sustained negativity in the violation condition (Royle & Courteau, 2014). To compensate for this problem, many studies adopted a post-target onset baseline (0–100 ms, Friederici et al., 1996; Hahne & Friederici, 1999, 2002; Neville, Nicol, Barss, Forster, & Garrett, 1991). However, using this type of baseline does not rule-out potential artefacts (especially in auditory studies) and rather ensures the appearance of a component at 100 ms. Despite these concerns, studies have continued to adopt this standard paradigm (e.g., Nichols & Joanisse, 2019; Sammler et al., 2013) which has serious potential to affect baselines.

Friederici's model emphasizes that her Phase I does not simply reflect the earlier temporal availability of syntactic category information: it also predicts that this information is *obligatorily* integrated first, and that it reliably guides later phases of language processing, such as the lexical-semantic one. Thus, without a grammatical syntactic representation that includes the current target word, Stage II, which includes semantic and morphosyntactic processing, cannot proceed. According to Friederici and collaborators, this claim is strongly supported by the absence of an N400 for combined syntactic and semantic violations (e.g. *Die Wolke wurde am *<u>geb</u>ügelt*, 'The cloud was on-the *<u>ironed</u>'), a phenomenon known as "semantic blocking". Semantic blocking is viewed as strong evidence in favour of the syntax-first model, independent of the ELAN. In fact, semantic blocking has been found even in studies that did not observe an ELAN effect (Frisch, Hahne, & Friederici, 2004).

However, it is crucial to note that as S&D point out, the standard paradigm used by Friederici and collaborators fails to create outright syntactic or semantic violations at the time when the target word is presented. First, contrary to what has been claimed, the prefix *ge-* is not a reliable marker for verb syntactic category: most past participles in German can be used as adjectival modifiers, and thus can appear after the preposition+determiner *am* 'on-the' (see S&D). Therefore, the occurrence of an ELAN at the prefix *ge-* cannot be associated with a failure in phrase-structure building due to a SC violation, because there isn't any. Following that logic, there cannot be any lexical-semantic incongruency at the target word either, since *<u>geb</u>ügelt* should be integrated as an adjective modifying a subsequent noun that has not yet appeared (and not the preceding subject NP *the cloud*). The participant must wait for the following noun to integrate lexical-semantic information, thereby explaining the absence of an N400 at the target word.

Just like the ELAN, the semantic blocking effect has not always been replicated (Luo, Zhang, Feng, & Zhou, 2010; Nickels, 2016; Van den Brink & Hagoort, 2004; Yang et al., 2015; Zhang et al., 2013). For example, in some studies, N400s have been observed in response to syntactic category violations when the participants were instructed to ignore syntax (Hahne & Friederici, 2002). Perhaps the strongest argument against semantic blocking comes from a replication of the original semantic-blocking study (Friederici et al., 1999) using the exact same German sentence materials, and which tested seven different groups of German native speakers while varying task instructions, linguistic profiles, and the types of filler sentences used, in order to elicit the ELAN and semantic-blocking effects (Nickels, 2016). Across all groups, a significant N400 was not only observed for combined syntactic and semantic errors, but even for pure syntactic category violations, suggesting that not only was the N400 not blocked, but that pure syntactic violations alone *can also* elicit N400s, at least when realized on a content word. Note however that the N400 for the pure semantic violation condition was the one with the largest amplitude (similar to Friederici et al., 1999), followed by the combined and pure syntactic anomalies. In other words, there was no additive effects of semantic and syntactic N400s on the combined anomaly condition. The slightly larger N400 in response to the combined violation compared to the pure syntactic anomaly could be interpreted as priming effects between the subject and the target word (Steinhauer & Drury, 2012). On the other hand, the uncertainty on how the target word should be categorized (adjective to be integrated with the upcoming noun or verb to be integrated with the previous subject) may explain why the N400 is smaller in the combined condition compared to the pure semantic one. Two other studies in Mandarin Chinese adopted a target-manipulation approach on the verb, rather than using a pre-verbal context-manipulation one (i.e., the standard paradigm), to investigate semantic blocking effects in the absence of morpho-syntactic cues for syntactic categories (Luo et al., 2010; Zhang et al., 2013). In object – subject – verb sentences the target verb would either be correct, semantically anomalous, or a combined semantic and syntactic category anomaly (that is a noun instead of a verb: e.g. Real estate business - corporation - recent several years - develop/*condition$_{[NOUN]}$ – […] 'This corporation has developed/*condition$_{[NOUN]}$ its real estate business […] during several recent years'), an N400 was observed in the combined condition, and therefore no semantic blocking was evidenced (Zhang et al., 2013). Similar results have been obtained using passive sentences and systematically manipulating for syntactic categories and semantic incongruency in Mandarin, where an N400 effect was also observed for "pure" syntactic category violations (e.g. 'That piece of glass is carefully wiped / *dishcloth […]', Yang et al., 2015). However, in these two studies, target nouns were always associated with a syntactic violation condition while target verbs were always correct, resulting in unbalanced designs that may affect the ERPs independent of syntactic category processing.

What are we left with? One the one hand, there is evidence in favour of an elegant, but very possibly incorrect model of sentence processing, and on the other, there is evidence pointing against it, but it is not consistent. In some cases, LAN effects are elicited by syntactic category errors, suggesting that LANs may reflect grammatical relations at large, including agreement and syntactic category violation effects (e.g., Steinhauer et al., 2009). In other cases, syntactic category violations elicit an N400 (Luo et al., 2010; Nickels et al., 2014; Yang et al., 2015; Zhang et al., 2013). If S&D are right and there is no reliable early marker for syntactic category identification, then a model for sentence processing should account for discrepancies between LAN and N400 findings. Zhang and colleagues (2013) offer a cross-linguistic explanation for the presence of an N400 in response to syntactic category violations in their studies on Mandarin. They suggest that, in the absence of morpho-syntactic cues, speakers rely more heavily on lexical-semantics for sentence processing, thus inducing an N400 in ungrammatical structures, while the existence of such cues would promote morpho-syntactic processes, thus eliciting an ELAN or a LAN, as found in German. However, N400s have also been evidenced in response to syntactic category violations with no semantic incongruency in German (Nickels, 2016), which makes this cross-linguistic explanation less compelling.

The interpretation of LAN effect as a defining marker of morpho-syntactic processing of local relationships (agreement in particular: Bornkessel & Schlesewsky, 2006; Friederici, 2002) has been discussed by Molinaro and colleagues (2011) on the one hand, and Tanner (2015) on the other. In a series of studies specifically looking at individual ERP patterns in response to agreement violations, Tanner and colleagues (2014, 2013; Tanner & Van Hell, 2014) provided evidence that, in their dataset, the LAN effect was largely the product of grand-averaging over multiple subjects. That is, the apparent LAN in grand average data resulted from component overlap between central N400s and right-posterior P600s found in the individual data sets. These two overlapping components cancelled each other out at central and posterior electrodes near the midline, and only left-lateralized portions of the N400 survived, thus resembling a LAN in the group average. Moreover, when plotting individual N400 magnitudes against P600 ones, Tanner et al. (2014) observed a negative correlation between these two measures. This correlation revealed that most individual ERP profiles did not display a biphasic (negativity+P600) profile (only 2/42 did), but instead tended either toward a P600 profile (31/42 participants) or an N400 profile (observed in only 9/42 participants). The authors suggest that the N400 may index mismatches with both word- and morphological-form based predictions in English speakers. Another approach does take the LAN effect at face value (Molinaro et al., 2011; 2015), while adopting a similar predictive approach to the one suggested by Tanner. Molinaro and colleagues (2011) propose that the LAN and N400 are at two points on a continuum reflecting a mismatch with predicted features: the more participants rely on lexical-semantic features, the more N400-like their response, and the more they focus on morpho-syntactic properties to process linguistic information, the more likely they will elicit a LAN. Their arguments are supported by an 80-participant analysis by Caffarra and colleagues, that does observe genuine LAN effects in response to agreement violations in 55% of participants (25% show an N400 effect, Caffarra, Mendoza, & Davidson, 2019). This discussion largely arose from studies of morpho-syntactic agreement (subject-verb, gender and number agreement, and tense marking), but could be extended to phrase-structure building and the identification of syntactic categories. As described above, many studies that do not employ Friederici's standard paradigm observe either a LAN (Hagoort, 2003; Steinhauer et al., 2009; Van den Brink & Hagoort, 2004) or an N400 (Luo et al., 2010; Yang et al., 2015; Zhang et al., 2013) rather than an ELAN for syntactic category violations.

When Tanner and colleagues observed distinct individual ERP profiles expressed as an N400 or a P600 dominance, they introduced the *Response Dominance Index* (RDI, Tanner & Van Hell, 2014) as a new measure that reflects the relative dominance of either the N400 (negative values) or the P600 effect (positive values for the RDI). They observed that individual RDIs were correlated across experimental manipulations (e.g. subject-verb or tense agreement). In other words, if an individual showed an N400 dominance for subject-verb agreement errors, this participant would likely also be N400 dominant in response to tense-agreement errors. It follows that some individuals might rely more on these lexically-based mechanisms (reflected by negativities), while others rely on combinatorial processes (indexed by the P600 effect) to process agreement, a pattern also found in response to syntactic category violations (Nickels, 2016).

An important methodological issue related to the response dominance index measure and especially with correlations between N400 and P600 amplitudes has to do with component overlap. In all previous studies, quantification of amplitudes for both components was done (i) in the same (centro-parietal) region of interest and (ii) using adjacent time-windows, i.e., the P600 window (e.g., 500-900 ms) directly followed that of the N400 (e.g., 300-500 ms). Since individual ERP components vary in latency and do not abruptly change at 500 ms, these correlations may at least to some extent result from component overlap – that is, early parts of P600s contaminating the N400 time-window and late parts of the N400 affecting the P600 interval (Brouwer & Crocker, 2017). These points are important because a systematic interdependence between N400 and P600 effects would also call for a reconsideration of most neurocognitive models of sentence comprehension that posit a biphasic detection-reanalysis pattern (Tanner & Van Hell, 2014) and assume that the N400 and the P600 reflect fundamentally distinct – and largely independent – cognitive processes.

The first objective of the present study is to re-evaluate the syntax-first approach to sentence processing by implementing a balanced design that controls both for contexts and targets. Second, this study aims to shed light on the nature of LAN vs. N400 negativities as indices of morpho-syntactic and lexical-semantic processing. Third, our data will allow us to test the (in)dependence between processes indexed by those negativities and the P600, while considering inter-individual variability.

1.1. The current study

This study uses an improved, novel paradigm that systematically manipulates contextual semantic priming and syntactic category of the target. The general idea was to create word category violations by replacing a noun with a verb (and vice versa), while ensuring that the context words immediately preceding the target noun or verb were optimally matched, in order to avoid any ERP baseline issues that rendered previous studies invalid (S&D). Nouns and verbs were selected as targets of the SC violation to maximize the comparability of our data with previous studies, but also because content words allowed us to include some additional manipulations that would specifically test the semantic blocking hypothesis (Friederici, Gunter, Hahne, & Mauth, 2004c; Friederici et al., 1999 vs. S&D). We selected two types of French sentence structures as carriers for our two target words. One carrier sentence (see example [13] below, experimental sentence in **bold font**) included a control verb (such as *oser* 'to dare') that requires an infinitive verb phrase as its complement (e.g., our target verb *plaquer* 'to tackle' in example [13]). The other carrier phrase (14) included a transitive verb (such as *ôter* 'to-remove') which mandatorily requires a noun phrase as its complement (e.g., *le* <u>*crapaud*</u> 'the <u>toad</u>'). All target nouns were directly preceded by a definite determiner (here: *le* 'the$_{[masc]}$'), the presentation interval of which would later be used to calculate the pre-target baseline for the ERP analysis. In order to ensure a comparable baseline in the verb condition (13), the target verb (*plaquer* 'to tackle') was preceded by a clitic pronoun (*le* 'him') that served as the target verb's direct object and, importantly, was a homograph of the determiner in the noun condition (i.e., *le plaquer* 'to tackle him'). Thus, conditions (13) and (14) included target nouns and target verbs in a grammatical sentence, and both target words were immediately preceded by the exact same written stimulus (here: *le*). In both experimental sentences, the target words were followed by the same 3-word prepositional phrase (here: *sur le côté* 'to/on the side') that would address spill-over effects and delay sentence final (wrap-up) effects in the ERPs (see S&D2012 for details). Finally, in order to introduce an antecedent for the pronouns and to license the use of definite determiners, we included context sentences (shown in normal font below) that preceded the experimental sentences. These context sentences also contained a semantic prime (shown in UPPER CASE font) for the respective target word in the experimental sentence (e.g., SWAMP is a prime for 'toad' in example [14]).

(13) *Marie et Jeanne jouent au HOCKEY avec leur copain.* ***Elles osent le <u>plaquer</u> sur le côté.***

Mary and Jane are playing HOCKEY with their friend$_{[M.SG]}$. **They$_{[F.PL]}$ dare him$_{[M..SG]}$ <u>tackle</u> to the side.**

'Mary and Jane are playing HOCKEY with their friend. They dare to <u>tackle</u> him to the side.'

(14) *Marie et Jeanne vont au MARAIS avec leur copain.* ***Elles ôtent le <u>crapaud</u> sur le côté.***

Mary and Jane go to the SWAMP with their friend$_{[M.SG]}$. **They$_{[F.PL]}$ remove the$_{[M.SG]}$ <u>toad</u> on the side.**

'Mary and Jane go to the SWAMP with their friend. They remove the <u>toad</u> on the side.'

The full experimental design with example stimuli is illustrated in Table V. We introduced outright syntactic category violations by swapping the target words (e.g., 'they remove the *<u>to-tackle</u>', second row in Table V), and manipulated semantic priming by interchanging context sentences so the target remains unprimed (e.g., prime: HOCKEY; target: <u>toad</u>, third row in Table V). Note that this manipulation also introduces lexical-semantic anomalies, so we will refer to unprimed sentences as *semantically anomalous*. In total, four main conditions involving both noun and verb targets were created by manipulating these two dimensions: syntactic category (Syntax: correct/incorrect), and lexical-semantic anomalies (Semantics: primed/unprimed). These were combined using a Latin-square design. Sentence contexts and targets appear in each condition, and all sentence pairs were matched on relevant psycholinguistic factors, presented in §2.2.

Table V. Sample stimuli for the eight experimental sub-conditions.

SYN	SEM		Context sentence	Experimental sentence: Subject + Verb	Clitic / Determiner	Target	Prepositional phrase
✓	✓	1	_Marie et Jeanne jouent au hockey avec leur copain._	_Elles osent_	_le_	_plaquer_	_sur le côté._
			Mary and Jane are playing hockey with their friend$_{MASC}$.	They$_{FEM}$ dare	him	tackle	on the side.
		2	_Marie et Jeanne vont au marais avec leur copain._	_Elles ôtent_	_le_	_crapaud_	_sur le côté._
			Mary and Jane are going to the swamp with their friend$_{MASC}$.	They$_{FEM}$ remove	the$_{MASC}$	toad$_{MASC}$	on the side.
✗	✓	3	_Marie et Jeanne vont au marais avec leur copain._	_Elles osent_	_le_	_*crapaud_	_sur le côté._
		4	_Marie et Jeanne jouent au hockey avec leur copain._	_Elles ôtent_	_le_	_*plaquer_	_sur le côté._
✓	✗	5	_Marie et Jeanne vont au marais avec leur copain._	_Elles osent_	_le_	_?plaquer_	_sur le côté._
		6	_Marie et Jeanne jouent au hockey avec leur copain._	_Elles ôtent_	_le_	_?crapaud_	_sur le côté._
✗	✗	7	_Marie et Jeanne jouent au hockey avec leur copain._	_Elles osent_	_le_	_*crapaud_	_sur le côté._
		8	_Marie et Jeanne vont au marais avec leur copain._	_Elles ôtent_	_le_	_*plaquer_	_sur le côté._

Hypotheses. In the 100-300 ms time-window of the target word, Friederici's syntax-first approach would predict an ELAN effect in ERPs for all syntactic category violations (i.e., Syntax: incorrect, rows 2 and 4 in Table 1) compared to the correct conditions. Following S&D, and because of our balanced design, we predict no difference between syntactically correct sentences and those with syntactic category violations in this early time-window. Analyzing the 100-300 ms window will allow us to test these predictions. Regarding the 300-500 ms time-window, several predictions can be made. First, Friederici's model would predict an N400 effect for lexical-semantic anomalies in otherwise grammatical sentences (Semantics: unprimed, Syntax: correct, line 2), but *not* for the combined violations due to semantic blocking (line 4). In contrast, we predict an N400 effect in *both* these conditions, as syntactic category violations are not expected to block lexical-semantic processing. Second, following Molinaro and colleagues (2011), since French has morpho-syntactic markers for syntactic category (e.g., infinitive verbs end with *–er*, *–ir*, or *–re*), one would expect syntactic category violations to elicit LAN effects, reflecting participants' use of these morpho-syntactic markers to detect anomalies. Tanner, on the other hand, would predict either an N400 or the "illusion" of a LAN effect in group data, due to component overlap of N400 and P600 components across participants (Tanner, 2015). Third, if the cognitive processes underlying syntactic category and lexical-semantic processing are modular (and thus independent), we predict additive effects of the syntactic and semantic manipulations in this time-window (Chow et al., 2014). On the other hand, if these processes rely on (and compete for) the same neurocognitive resources, we predict an interaction of these two factors on the ERP responses (Hagoort, 2003a). After 500 ms, most – if not all – frameworks would predict P600 effects at least for all conditions involving syntactic category violations (Syntax: incorrect, lines 2 and 4). According to some studies, P600 modulations may also be seen for semantic manipulations (e.g., Steinhauer et al., 2010). If so, similar predictions regarding additivity versus interaction between syntactic and semantic effects apply as for the 300-400 ms time-window. We will assess these predictions while taking inter-individual variability into account, following Molinaro and colleagues' recommendations for analyzing the data using mixed-effect models (Molinaro et al., 2015), and further explore our individual datasets following Tanner et al.'s practice (Tanner et al., 2014, 2013; Tanner & Van Hell, 2014). The latter analyses should reveal if the amplitudes of P600s and preceding negativities are correlated, and whether such correlations are partly due to component overlap. In other words, our data are expected to not only clarify certain predictions of Friederici's (2002, 2012) influential model, but also shed light on a number of methodological issues in recent ERP research.

2. Methods

2.1. Participants

Thirty-nine native speakers of Quebec French (20 of which identified as women) aged 20–31 (mean age: 25;10) participated in the study. All were right-handed (confirmed with an abbreviated French version of the Edinburg handedness questionnaire (Oldfield, 1971), had normal or corrected-to-normal vision, and reported no neurological disorder. Given the highly multilingual environment in Montreal, participants filled in a language usage questionnaire where they were asked to evaluate their daily exposure to French at work or school, in the family, social circles, and through media use. All of the participants had limited exposure to English (*Mean % daily exposure* = 11.41, SD = 10.69), and reported having learned English as part of their education program only (*Mean age of first exposure* = 12.44, SD = 1.94). Additional native French speakers were recruited to evaluate our stimuli in three separate offline experiments: an acceptability rating study (*n* = 40), a relatedness rating questionnaire (*n* = 68), and a cloze test (*n* = 40).

2.2. Stimuli

In total, 160 pairs of correct sentences as in (13). and (14). were created. In order to minimize inter-item variability, we controlled for the three following dimensions: (1) phonological and lexical properties of context and target words, (2) acceptability ratings for correct and anomalous conditions, and (3) degree of priming between primes and target words. Being aware that phonological structure may result in differences in the early ERP time-windows (that may then be misinterpreted as ELANs; see S&D), and that lexical properties can affect the N400 component, we matched (i) control verbs with transitive verbs in the carrier sentences, and (ii) target verbs with nouns, on both phonological/orthographic and lexical dimensions (see Chapter 1 for details about the stimuli). Second, any difference in acceptability within pairs in the correct or the anomalous conditions could also result in unwanted ERP differences. Focusing on correct sentences and combined anomalies, we distributed all 160 sentence pairs into four experimental lists, using a Latin square design, so that every participant would only see one sentence per pair and condition. Forty undergraduate students were asked to rate their acceptability on a scale from 1– *acceptable* to 4– *unacceptable*. Results showed that, although our correct sentences were judged more acceptable than our anomalous sentences ($p < .001$), there was no effect of target lexical category (noun or verb, $p = .514$) nor interaction between these two factors (acceptability*lexical category: $p = .196$). We thus ensured that no major lexical difference within sentence pairs could affect acceptability judgements or electrophysiological responses during the online experiment. Third, we wanted to control for priming effects on target nouns and verbs, as more priming would result in a reduced N400 in the correct condition. We thus sought to match target nouns and verbs in their degree of priming. A questionnaire assessing relatedness between primes and targets (from 1: not related to 5: extremely related) revealed no difference between noun and verb targets, despite highly related prime-target pairs (for all conditions $M = 3.35$, $SD = 1.07$, $p > .1$). However, we also used a cloze test to evaluate the predictability of target words given the context sentence, and computed a cloze probability index ranging from 0 (never predicted by our participants) to 1 (always predicted). Paired t-tests revealed that verb targets (*Mean* = .1) were less predictable than nouns (*Mean* = .17, $p = .006$) despite a general low predictability level. Priming results and methods are summarized in Chapter 1 §3.3). Since it was impossible to control for this specific dimension without interfering with other lexical, formal, acceptability, and relatedness dimensions that we had carefully controlled for, we opted to include cloze probability as a random slope in our statistical models (see §3.2). It is important to note that the experiment is designed so that results on target nouns and verbs will be merged to perform the analyses: this is key to avoid context effects that have previously led to baseline issues, and so that possible lexical differences on target words would not play any role.

Creating the various violation conditions based on the original set of 160 grammatical sentence pairs resulted in a final set of 1280 items, with an average of 13.24 words per item ($SD = 1.12$; range for context sentences: 5-10 words; number of words for target sentences = 7), corresponding to an average duration of 6622 ms per trial ($SD = 560$ ms). The 1280 items were divided into four lists using a Latin square design, such that each participant would read one single list with 320 sentences (80 per condition), but no prime or target word would ever be repeated within a given list. To every list, we added the same set of 80 filler sentences that were either correct or contained one or more subject-verb number-agreement error(s) unrelated to the present study, such that each participant read a total of 400 experimental sentences with their corresponding contexts.

2.3. Experimental procedure

Participants sat in a chair 80 cm in front of a computer monitor and read the sentences while their electroencephalogram (EEG) was recorded. Each trial began with the presentation of a fixation cross (500 ms). The sentences were presented in rapid-serial-visual presentation mode (each word: 300 ms display + 200 ms blank screen), and words appeared in white 30-point Arial font on a black background. At the end of every trial, a "???" prompt appeared and remained on a screen until participants scored sentence acceptability on a scale of 1 to 5 (1: *totally acceptable* – 5: *totally unacceptable*) by pressing a button on a response box. After participants responded, a second prompt "!!!" would appear for a 1800 ms interval during which participants were encouraged to blink their eyes (in order to avoid eye blinks during sentence presentation). The experimental sessions started with two short practice blocks with six sentences each, while the 400 experimental trials were divided into eight blocks of 50 trials each, separated by short breaks. The recording lasted about 2.5 hours including setup. Consent forms and all materials were approved by the Ethics Review Board of the Faculty of Medicine, McGill University, and *Comité d'éthique de la recherche en santé, Université de Montréal*.

2.4. EEG recording and data processing

EEG was recorded continuously from 21 Ag-Cl active-shield electrodes mounted on an EEG cap (Waveguard™ original, *ANT Neuro*, Netherlands) according to the 10-20 system (Jasper, 1958) at the following sites: FP1-FPZ-FP2-F7-F3-FZ-F4-F8-T3-C3-CZ-C4-T4-T5-P3-PZ-P4-T6-O1-OZ-O2, with a 512 Hz sampling rate and a 0.001–100 Hz online forward filter.

All EEG electrodes were referenced online against the right mastoid. An electrode between FPZ and FZ served as ground. Impedances were kept below 5 kΩ.

Data were analyzed using *EEGLab* (Delorme & Makeig, 2004) and *ERPLab* (Lopez-Calderon & Luck, 2014). Continuous data were re-referenced offline to linked mastoids, and bandpass-filtered with .1 and 40 Hz cut-off frequencies (IIR Butterworth filter). After epoching the data from -200 to 2000 ms relative to the onset of the target word, we rejected data that exceeded a peak-to-peak threshold of 75 μV (in 100 ms steps). Three participants were excluded due to excessive artefacts (over 50% in at least one of the eight sub-conditions), and one for not respecting instructions. The remaining 36 participants had 11.5% rejected trials on average, with no differences among the eight sub-conditions, $F(7,217) = 1.836$, $p =.101$, Greenhouse-Geisser corrected. Finally, the epoched data was corrected using a 200 ms pre-target baseline. Based on previous findings and visual inspection of our own ERP data, we selected four analysis time-windows corresponding to the following components: ELAN (100-300 ms), N400 (350-500 ms), early P600 (550-650 ms), and late P600 (800-1200 ms).

2.5. Statistical analyses

All statistical analyses were done using *R* version 3.4.2 (*Short summer*, R Core Team, 2017). Since we are interested in whether syntactic category identification interacts with lexical-semantic processing at both performance and electrophysiological levels, we implemented a Syntax×Semantics (2×2) design in both behavioural and ERP analyses.

2.5.1. Behavioral data analysis using cumulative link mixed effects models

Acceptability ratings on a Likert scale (from 1: totally acceptable to 5: totally inacceptable) were analyzed using cumulative link mixed effects models using the *clmm* function from the *ordinal* package (Christensen, 2015). Cumulative link models are more appropriate than parametric statistical tests (e.g., linear regressions) when analyzing ordinal data such as Likert scales (Bauer & Sterba, 2011). Considering the ratings as continuous variables in statistical tests can be problematic. Because Likert scales are limited on the edges, they lead to a "censoring" effect (in our case, participants cannot select 0 or 6, Poschmann & Wagner, 2016). This limitation brings the mean closer to the center and decreases variability in the data, both of which can affect the statistical test. Second, points on the scale are not always processed as continuous by participants. Cumulative links models assume that the ratings from 1 to 5 are ordered, but not that the five points on the scale are equidistant, or that the values beyond these points are interpretable. For a given factor, the estimate expresses the probability of falling above or below the baseline level (here, the correct condition). A positive estimate therefore expresses a higher rating on the scale. To ensure convergence, the random effect structure was limited to random slopes for condition per participant. We fit a model that included Syntax, Semantics (two levels, correct and anomalous), and their interaction.

2.5.2. ERP data analysis using mixed effects models

ERP analyses were run using mixed effect models (packages: *lme4* and *lmerTest* (Bates, Mächler, Bolker, & Walker, 2015; Kuznetsova, Brockhoff, & Christensen, 2017), because they can adjust for repeated measures of both participants and items while controlling for individual variability, and handle missing data and unequal sample sizes better than traditional ANOVAs (Gelman & Hill, 2007). ERP effects on the midline and lateral sites were analyzed separately. The maximal random structure ensuring convergence included random slopes for condition per participant, condition per item, and cloze probability per item. We first calculated main effects and interactions for factors Syntax, Semantics, Anteriority (two levels: using F and C electrodes as anterior sites; and P and O electrodes as posterior sites), and Hemisphere (two levels: left and right). Then, we decrementally removed interactions and factors from this full model until we reached the optimal model, determined by comparing two minimally different models using ANOVAs. For clarity of interpretation, we used ANOVA wrappers (Type III Wald chi-square test) with the *car* package (Fox & Weisberg, 2011). When needed, we performed follow-up analyses of interactions and post-hoc pairwise comparisons using the *emmeans* package (Lenth, 2018).

3. Results

3.1. Behavioral data

Sentences with either a lexical-semantic anomaly or a syntactic category violation were judged to be less acceptable than correct sentences (M = 2.02, SD = 0.47), as supported by the strong main effects of Syntax and Semantics in Table VI. Further, the interaction between Syntax and Semantics and the follow-up analyses show that lexical-semantic anomalies primarily affected judgments in the absence of a syntactic anomaly (M = 3.09, SD = 0.61). In contrast, sentences containing a syntactic category violation were judged equally unacceptable regardless of whether lexical-semantic anomalies were present (M = 3.93, SD = 0.67) or not (M = 3.71, SD = 0.73).

Table VI. Effects of syntactic category violations and semantic anomalies on participants' responses.
Cumulative Link Mixed Model fitted with the Laplace approximation.

		Effects	Estimate*	SE	z-value	p-value
Main model		Syntax	3.099	0.251	12.34	< 0.001
		Semantics	1.971	0.179	10.98	< 0.001
		Syntax×Semantics	-1.525	0.155	-9.84	< 0.001
Follow-up pairwise comparisons	Semantics correct	Syntax	1.508	0.027	56.13	< 0.001
	Semantics anomalous	Syntax	0.624	0.031	20.34	< 0.001
	Syntax correct	Semantics	0.938	0.027	34.60	< 0.001
	Syntax incorrect	Semantics	0.054	0.028	1.91	< 0.06

Note: in the main model, the estimate expresses the probability of falling above or below the rating for the correct condition.

3.2. ERP data

The ERP waves for the correct condition and each of the three anomalous conditions (syntactic category violations, lexical-semantic anomalies, and combined anomalies) are illustrated in Figures 3 and 4. All ERP waves are time-locked to target word onsets: verb and noun targets are merged, thus canceling out any possible context or target effect. Visually, two observations are striking. First, we observe the absence of the ELAN: our balanced design and the care with which we selected and controlled the stimulus materials led to virtually perfectly aligned ERP onsets after using a 200 ms pre-target baseline correction. Note that our pre-stimulus baselines themselves are also perfectly aligned across all conditions, contrary to most syntactic category violation studies that use a "post target onset" baseline (0–100 ms) due to unbalanced designs (Lenth, 2018). Second, we observe an N400 in both syntactic category violation conditions. We will investigate these observations by running statistical analyses on the 100-300 ms time-window (where one would predict an ELAN effect) and the 350-500 ms time-window for the N400 effect. Later effects are also observable: a posterior positivity in both the syntactic category violation and the 'combined' conditions between 500 and 1200 ms, plus a frontal positivity between 500 and 650 ms in the pure syntactic category violation condition only. We ran mixed-effect models on each of these time-windows and report significant main effects and interactions involving factors Syntax or Semantics in Table VII. These effects will be commented in text, along with follow-up analyses when necessary.

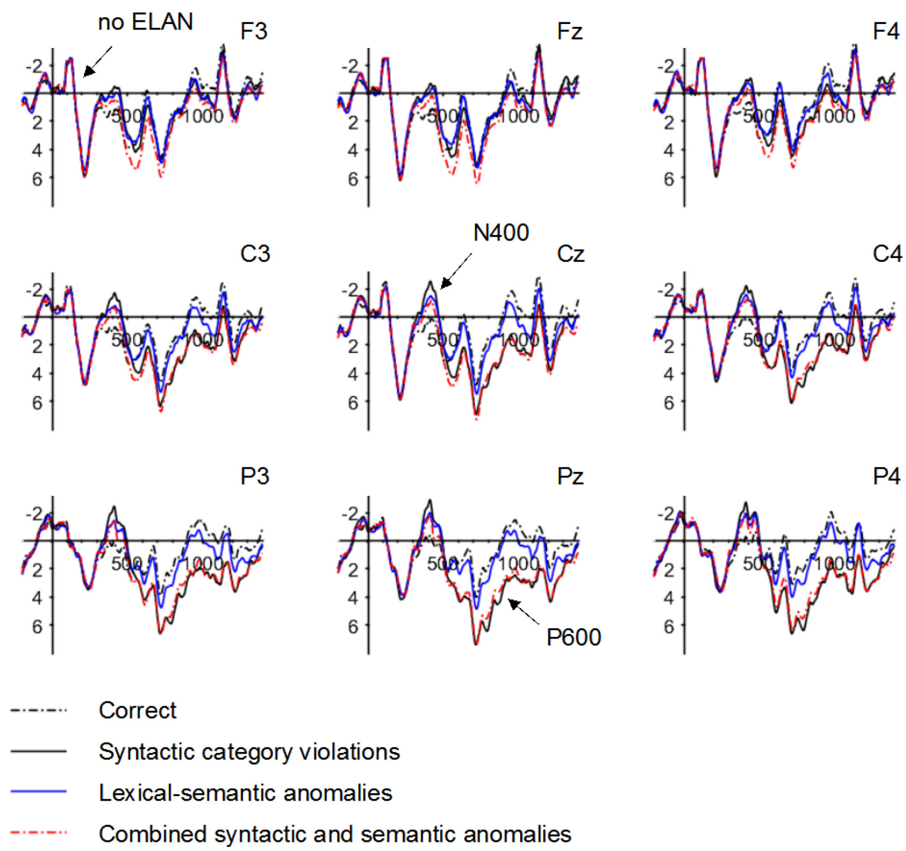


Figure 3. Grand average waveforms for the four experimental conditions. ERPs are shown time-locked to target word onset for correct sentences (dotted black line), sentences with a syntactic category violation (solid black line), sentences with a semantic anomaly (blue line), and sentences with combined anomalies (dotted red line) on nine representative electrodes. Target onset is indicated by the vertical bar, where tick bars represent 2 µV of activity; time-windows extend from -200 ms to 1300 ms relative to target onset.

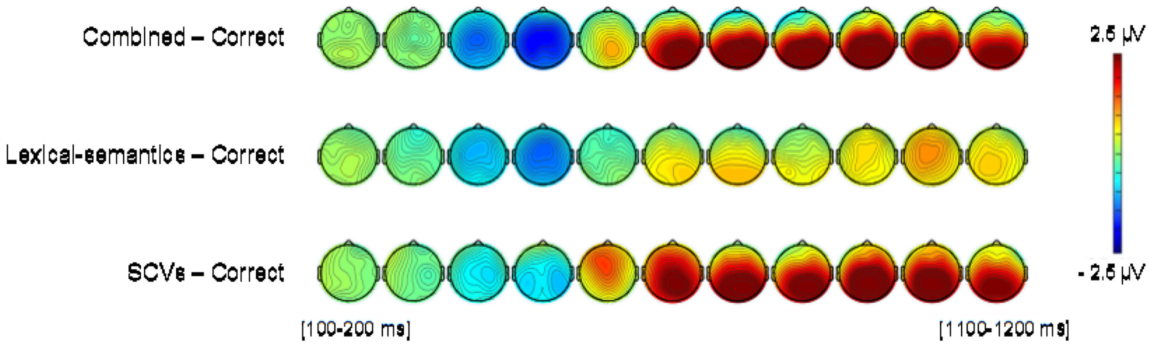


Figure 4. Voltage maps illustrating the effect of each of the three anomalous conditions. Mean amplitudes measured between 100 ms and 1200 ms in 100 ms increments.

Table VII. Analysis of deviance table (Type III Wald chi-square tests). Significant effects corresponding to the main mixed-effect models in the four time-windows of interest (ELAN, N400, early positivity, posterior P600), at midline electrodes and lateral sites.

Time-window	Site	Fixed effects and interactions	Chi-square	Df	p-value
100-300 ms	Midline	(Intercept)	65.097	1	< 0.001
	Lateral sites	(Intercept)	50.978	1	< 0.001
350-500 ms	Midline	(Intercept)	19.633	1	< 0.001
		Syntax	6.623	1	0.01
		Semantics	27.085	1	< 0.001
	Lateral sites	(Intercept)	34.180	1	< 0.001
		Syntax	8.546	1	0.003
		Semantics	22.890	1	< 0.001

Table VII (continued)					
Time-window	**Site**	**Fixed effects and interactions**	**Chi-square**	**Df**	**p-value**
550-650 ms	Midline	(Intercept)	27.403	1	< 0.001
		Syntax	21.971	1	< 0.001
		Syntax×Semantics×Anteriority	6.088	1	0.014
	Lateral	(Intercept)	31.970	1	< 0.001
		Syntax	12.705	1	< 0.001
		Syntax×Semantics×Anteriority	5.481	1	0.019
800-1200 ms	Midline	(Intercept)	0.229	1	0.632
		Syntax	16.544	1	< 0.001
		Syntax×Anteriority	13.554	1	< 0.001
		Syntax×Semantics×Anteriority	6.098	1	0.014
	Lateral	(Intercept)	1.327	1	0.249
		Syntax	9.226	1	0.002
		Semantics	4.227	1	0.040
		Syntax×Semantics	3.860	1	0.049
		Syntax×Anteriority	64.223	1	< 0.001
		Syntax×Hemisphere	4.763	1	0.029
		Syntax×Semantics×Anteriority	4.398	1	0.036

3.2.1. Absence of ELAN: 100-300 ms

Visual inspection of the data (Figures 3 and 4) indicates no early anterior negativity elicited by syntactic category violations, as the onset components in all four conditions are nicely aligned until around 300 ms. There was no statistically significant effect of Syntax, Semantics, and no interaction involving these terms in this early time-window.

3.2.2. Main effects of Syntax and Semantics on the N400: 350-500 ms

Visual inspection of the data in the 350-500 ms time-window revealed a centrally distributed negativity elicited by both syntactic category violations and lexical-semantic anomalies, while the combined syntactic and lexical-semantic anomaly condition elicited a larger N400 than the "pure" syntactic and semantic conditions. At both midline and lateral sites, the best model fit included main effects of Syntax and Semantics without any interaction involving these factors (Figure 5), suggesting that a broadly distributed N400 was elicited under both violation types, with an additive effect of syntactic category violations and lexical-semantic anomalies. We further tested if the N400 effects elicited by syntactic category violations and lexical-semantic anomalies have different scalp distributions by calculating the difference waves between each of these two conditions and the correct one, and running a new model with factors Condition (2 levels: lexical-semantic anomalies, syntactic category violations) and Electrode (21 levels). To calculate those difference waves, we had to average over items, so the random structure was highly simplified to random slopes for participants per condition, making the model must less conservative and maximizing the chance of getting significant effects. Even so, there was no effect of Condition and no interaction between Condition and Electrode. Finally, in order to better understand how priming effects modulate the N400 response in different conditions, we explored whether cloze probability had an impact on the N400 amplitude under syntactic and semantic manipulations. We selected central electrodes (C3-Cz-C4) and ran a mixed-effects models with factors Syntax, Semantics, and Cloze probability. Interestingly, we observed not only main effects of Syntax ($X^2(1) = 9.56$, $p = .002$), Semantics ($X^2(1) = 32.12$, $p < .001$), and Cloze ($X^2(1) = 5.69$, $p = .017$), but also an interaction between Cloze and Semantics ($X^2(1) = 6.8$, $p = .009$). Follow-up analyses using a regression tree from the *partykit* package, Hothorn & Zeileis, 2015) confirmed that while semantically-anomalous sentences induced a larger N400 amplitude than primed sentences (p < .001), cloze probability modulated the semantically primed sentences regardless of syntactic category violation. The N400 elicited by the target had a smaller (less negative) amplitude when the target was highly probable (cloze probability above .4, $p = 0.001$) compared to less probable sentences (all full models can be found in Appendix 5). Note however that this analysis is exploratory because we did not explicitly control for cloze probability, therefore the distribution of cloze probability values is largely skewed towards zero.

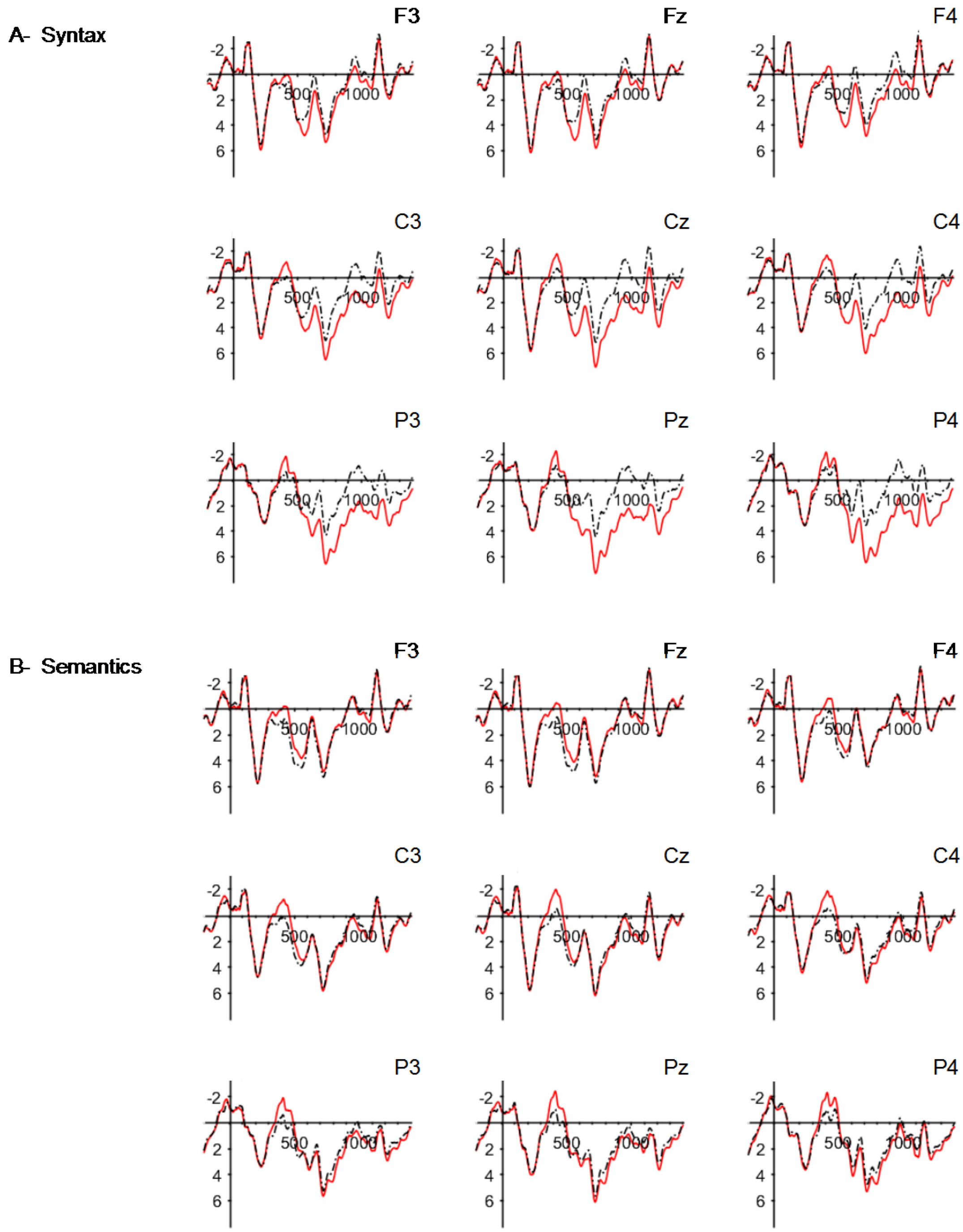


Figure 5. Grand average waveforms illustrating the main effects of A- Syntax and B- Semantics. Anomalous trials are in red and correct ones in black. Target onset is indicated by the vertical bar, where tick bars represent 2 μV of activity; ticks marks on the horizontal line represent 500 ms of time.

3.2.3. Syntactic category violation effects on the Early Positivity: 550-650 ms

Inspection of the four individual conditions suggests that violations that are uniquely syntactic elicit a positivity that appears at frontal sites as early as 550 ms and becomes posterior around 650 ms, while combined anomalies only elicit a posterior positivity. These observations were reflected by a main effect of Syntax at midline and lateral sites, and a Syntax×Semantics×Anteriority interaction. As follow-up analyses that focused on anterior and posterior sites did not converge, we split our data by levels of Syntax. While no significant effect of Semantics×Anteriority was found for syntactically correct sentences, the model focusing on syntactically anomalous sentences showed a Semantics×Anteriority interaction (midline: $X^2(1) = 14.05$, $p < .001$; lateral sites: $X^2(1) = 7.16$, $p = .007$). Pairwise comparisons suggest an anteriority effect for syntactic category violations, with Anterior sites being more positive (Anterior – Posterior = 0.73 μV, z = 3.73, $p < .001$). Visual inspection of Figure 4 indicates that this anterior effect is specific to pure syntactic category violations.

3.2.4. Late P600: 800-1200 ms

Two patterns are of interest when considering our results together with the ERP waves in Figure 3 and the voltage maps in Figure 4, namely (a) a large P600 effect for syntactic category violation (irrespective of semantic anomalies), and (b) a small P600 elicited by pure lexical-semantic anomalies. At midline sites, there was a main effect of Syntax ($p < .001$), and interactions between Syntax×Anteriority ($p < .001$) and Syntax×Semantics×Anteriority ($p = .015$). At Lateral sites, we observed a similar main effect of Syntax ($p = .002$), a main effect of Semantics ($p = .04$), and interactions of Syntax×Anteriority ($p < .001$), Syntax×Semantics ($p = .049$), Syntax×Semantics× Anteriority ($p = .036$), and Syntax×Hemisphere ($p = .029$). For each model, separate follow-up models for anterior and posterior sites confirmed that syntactic category violations elicited a large, mostly-posterior positivity (midline at anterior sites: $X^2(1) = 21.07$, $p < .001$; midline at posterior sites: $X^2(1) = 42.27$, $p < .001$; lateral electrodes at anterior sites: $X^2(1) = 10.51$, $p = .001$; lateral electrodes at posterior sites: $X^2(1) = 37.75$, $p < .001$). We further explored the interactions involving Syntax×Semantics by splitting into levels of Syntax (correct, incorrect). Main effects of Semantics were observed for syntactically-correct sentences, suggesting a small broadly-distributed positivity in response to lexical-semantic anomalies (midline: $X^2(1) = 3.69$, $p = .05$; lateral electrodes: $X^2(1) = 4.43$, $p = .03$). Significant interactions of Semantics×Anteriority were observed for syntactically incorrect sentences (midline: $X^2(1) = 8.25$, $p = .004$; lateral electrodes: $X^2(1) =10.65$, $p = .001$), potentially revealing a slightly larger posterior positivity for combined anomalies compared to pure syntactic category violations (syntactic category violations effect at midline: Anterior – Posterior = -0.98 μV, z = -5.67, $p < .001$; Combined anomalies effect at midline: Anterior – Posterior = -1.69 μV, z = -9.67, $p < .001$; syntactic category violations effect at lateral electrodes: Anterior – Posterior = -0.56 μV, z = -6.89, $p < .001$; Combined anomalies effect at lateral electrodes: Anterior – Posterior = -0.8 μV, z = -9.57, $p < .001$). Note however that comparing between pure syntactic category violations and combined anomalies at Anterior and Posterior levels did not reveal any significant differences.

To summarize our findings so far, both semantic and syntactic anomalies seem to have elicited N400s and P600s, with seemingly additive effects on the N400 and interactive effects on the P600. The next section will investigate whether these biphasic ERP profiles in the grand-average data were representative for the individual data.

3.3. Exploratory analyses of individual data

Our analyses revealed N400 effects (and no LAN effect) in response to both syntactic violations and lexical-semantic anomalies. The N400 was followed by a large P600 effect only for sentences with a syntactic category violation. Visual inspection of individual data, however, revealed some interindividual variability: as shown in previous studies, individuals seem to display different dominance toward either an N400 or a P600 profile (Nickels, 2016; Tanner & Van Hell, 2014). In this section, we evaluate whether (1) larger N400s equate to smaller P600s across individuals (i.e. whether the N400 and P600 are negatively correlated), and (2) whether their ERP responses are similar across conditions (i.e. whether N400s and P600s observed in response to one condition each correlated with N400 and P600 effects in *other* conditions).

3.3.1. Correlations between N400 and P600 effects within each condition

The magnitudes of the N400 and P600 effects were estimated for every individual by calculating the difference between each of the three anomalous conditions (SC, lexical-semantic, and combined anomalies) minus the correct condition and by quantifying the amplitudes of these difference wave in representative time intervals. We then calculated the correlation between N400 and P600 amplitudes within all three conditions (using *Hmisc*, Harrell, 2019). As we were concerned that component overlap may contribute to spurious correlations, both here and in previous studies (Brouwer & Crocker, 2017), we ran two analyses. The first one adopts the method promoted by Tanner and colleagues, i.e., using the same region of interest (C3-Cz-C4-P2-Pz-P4) to quantify both N400 and P600 amplitudes, and using time-windows that were almost adjacent (50 ms apart), i.e., 350–500 ms for the N400 and 550–650 ms for the (early) P600. As expected, in this analysis we observed significant negative correlations in the SC ($r = -.61$, $p < .001$) and lexical-semantics conditions ($r = -.55$, $p < .001$), and a marginal correlation in the combined condition ($r = -.32$, $p = .057$). The second analysis aimed to minimize component overlap and allowed us to test our hypothesis that component overlap may have inflated previous findings. A first difference to previous approaches was that the regions of interest were optimized and limited to electrodes where each component was most prominent: the N400 was measured at C2-Cz-C3 electrodes and the P600 at P2-Pz-P3. A second difference was that the time-windows were selected 300 ms apart (N400: 350–500 ms; P600: 800–1200 ms). As illustrated in Figure 4, a significant negative correlation was observed *only* in the lexical-semantic condition ($r = -.54$, $p < .001$), which remained relatively unchanged compared to the first analysis. For the syntactic conditions (which elicited much larger P600s) the absence of a correlation between N400 and P600 amplitudes seems to confirm that component overlap may have played a role in our first analysis (syntactic category violations: $r = -.3$, $p = .08$; combined: $r = -.09$, $p = 0.60$).

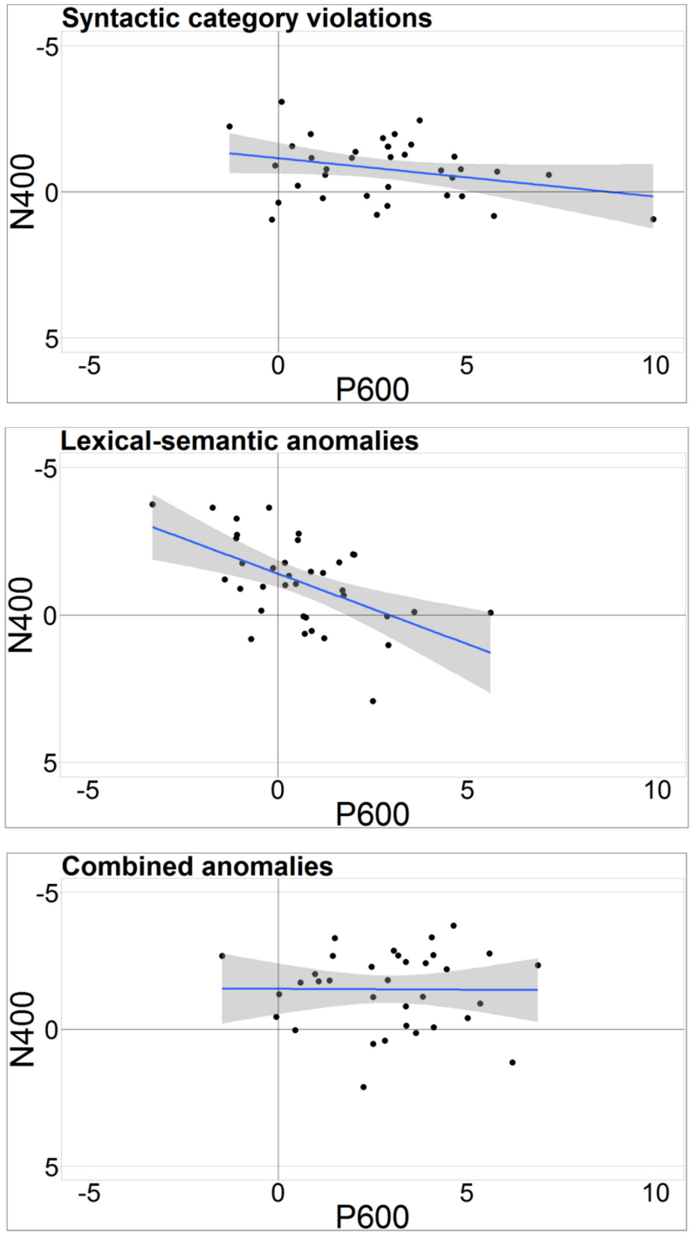


Figure 6. Scatterplots showing the relationship between N400 and P600 effects across individuals in the three anomalous conditions. For each participant, we subtracted the average response (in µV) to syntactic category violations, lexical-semantic anomalies, and combined anomalies minus the correct condition. Linear smooths are fitted with a 95% confidence interval.

Next, we addressed whether participants elicited similar responses throughout the experiment: for example, whether a participant who elicited a large N400 in responses to lexical-semantics anomalies also showed a large N400 effect in the two other conditions. This pattern could be taken as an indicator that ERP components largely reflect individual strategies independent of experimental manipulations. We therefore compared the response magnitudes of the N400 and P600 elicited across conditions (Nickels, 2016).

3.3.2. Correlations for N400 and P600 effect magnitudes across conditions

We calculated correlations for the N400 and the P600 effects between all possible pairs of our three violation conditions. When appropriate, we compared the correlation coefficients (*cocor*, Diedenhofen & Musch, 2015) using a test statistic that compares two correlation coefficients based on dependent groups with one overlapping variable (Silver, Hittner, & May, 2004). All results are illustrated in Figure 5. The N400 effects correlated positively when comparing SC and combined anomalies ($r = 0.44$, $p = 0.007$), as well as lexical-semantic and combined anomalies ($r = 0.48$, $p = 0.003$), but no significant N400 correlation was found when comparing SC and lexical-semantic anomalies ($r = 0.3$, $p = 0.076$). There was no difference between these coefficients, even when comparing the highest ($r = 0.48$) and the lowest one ($r = 0.3$; $z = -1.089$, $p = 0.276$), probably because the correlation was moderate even when significant. On the other hand, P600 effects were highly correlated across all conditions: SC and combined anomalies ($r = 0.83$, $p < 0.001$); lexical-semantic and combined anomalies ($r = 0.69$, $p < 0.001$); and SC and lexical-semantic anomalies ($r = 0.73$, $p < 0.001$).

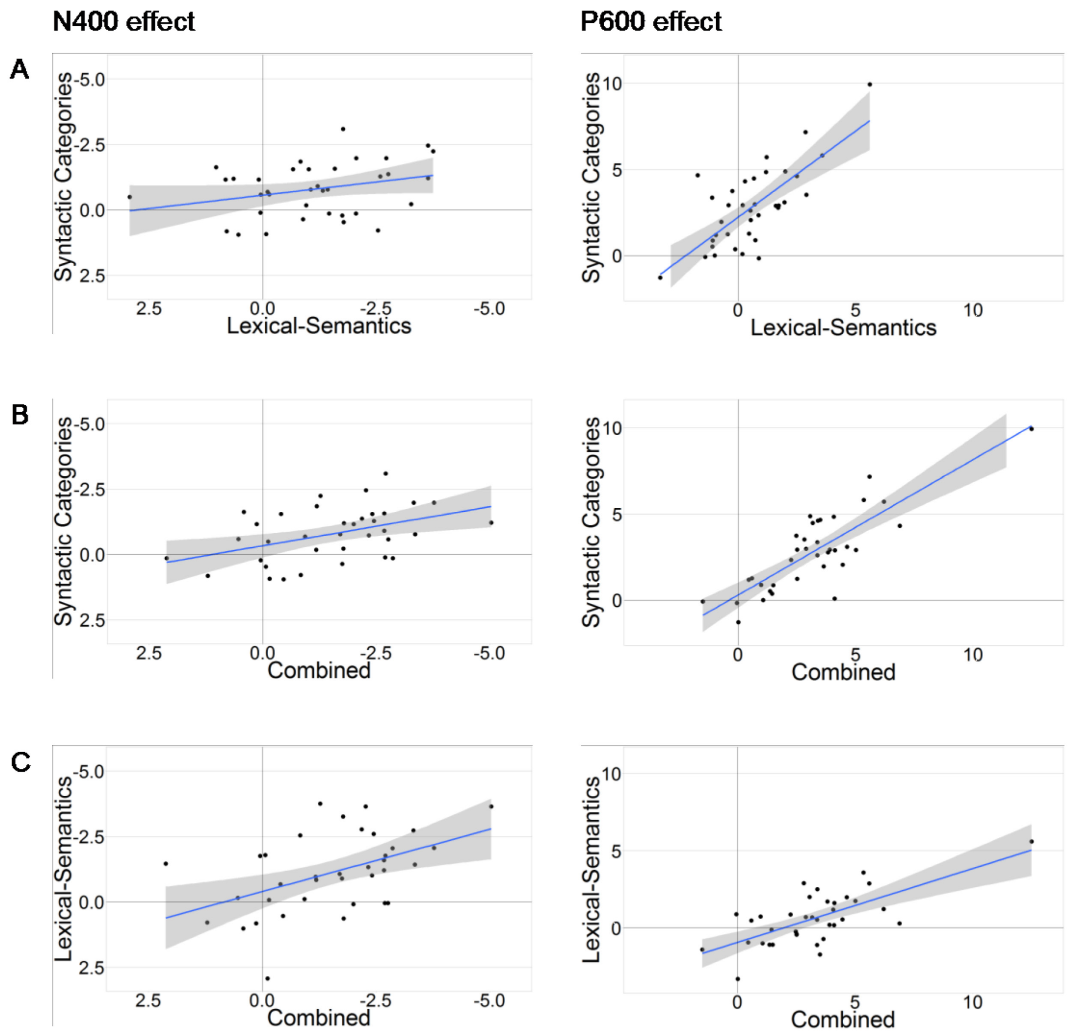


Figure 7. Scatterplot showing the relationship of N400 and P600 effects across individuals *between* conditions. We plotted the average response (in µV) for: A- syntactic category violations against lexical-semantic anomalies, B- syntactic category violations against combined anomalies, and C- lexical-semantic anomalies against combined anomalies. Linear smooths are fitted with a 95% confidence interval.

4. Discussion

4.1. Consequences for “syntax-first” approaches to sentence processing

The present study reevaluates the temporal organization of syntactic category and lexical-semantic processing, and demonstrates that syntactic category violations do not elicit early ERP responses (previously described as ELANs (Friederici, 1995, 2002; Friederici et al., 1993), nor do they block lexical-semantic information processing. To compare and contrast our findings with Friederici’s model, we summarize the time-course of ERP effects in response to syntactic category violations, lexical-semantic, and combined anomalies in Figure 8.

Figure 8. Difference waves between each of the anomalous conditions minus the correct condition. The effects of syntactic category violations (in red), lexical-semantic anomalies (in blue), and combined anomalies (in black) are illustrated on three representative electrodes: F3 (where no ELAN is found), Cz (for the N400), and Pz (for the P600). Target onset is indicated by the vertical bar, where tick bars represent 2 μV of activity; ticks marks on the horizontal line represent 500 ms of time.

Our study is one of the few studies that employ outright syntactic category violations (Bowden, Steinhauer, Sanz, & Ullman, 2013; Steinhauer et al., 2009). In studies employing the standard paradigm, syntactic category violation sentences could always be rescued by adding morphological or lexical information after the target (e.g., _Die Bluse wurde am *gebügelt_, ‘The blouse was on-the *ironed’ → _Die Bluse wurde am gebügelt**en** Jackett_ *befestigt,* 'The blouse was pinned to-the ironed jacket'). In our experiment, we use transitive or control verbs that specifically select for either a noun or a verb, respectively, such that swapping targets automatically results in outright syntactic category violations. We argue that our data are unambiguous in that they provide a strong argument against the primacy of syntactic categories (and, by extension, phrase-structure building), over other types of information, such as lexical-semantic processing.

It has been argued that the ELAN effect is a reliable marker of syntactic processing in auditory studies, but is not always expected in visual paradigms even when using the same linguistic materials (Bowden et al., 2013; Steinhauer et al., 2009). That is, the absence of an ELAN in a reading study can still be argued to be unsurprising as it has been claimed to be more strongly linked to auditory sentence processing (Friederici, 2002). However, this modality difference (i.e., the absence of ELANs in many visual studies) is said to be due to specific shortcomings in experimental designs. Thus, Gunter and Friederici (1999) suggested that the ELAN in reading studies can *only* be observed under certain "visual input conditions" that involve high contrast and a 300 ms stimulus-presentation time followed by a 200 ms inter-stimulus interval. Importantly, we used these exact specifications in our study and still did not observe an ELAN. Finally, one could also claim that the absence of ELANs in many reading studies simply indicates that there is no early ERP marker for syntactic category violations, and that the finding of ELAN components in many auditory studies is in reality an artefact due to modality-specific issues, for example prosodic differences between conditions (S&D2012). In order to test if our present findings are indeed modality-specific, we are currently following up with an auditory experiment.

Moreover, our data clearly show that an N400 is present in all conditions containing a syntactic category violation, therefore providing an important piece of evidence against semantic blocking, and directly contradicting the "syntax-first" model. Contrary to the standard paradigm used in many previous syntactic category violation studies (Friederici et al., 1996; Hahne & Friederici, 1999, 2002; Isel et al., 2007; Neville et al., 1991) , in our design, both SC information and lexical-semantic information are available and unambiguous at once (Dikker et al., 2009). Because the target word is in a complement position, anomalous sentences cannot be "fixed" after the target. Therefore, not only is our design one of the rare ones that employ outright syntactic category violations, but it is also one of the very few ones (possibly the first, as far as we know) to systematically control for true syntactic category violations together with lexical-semantic anomalies. This piece of evidence complements previous findings – such as a large-scale replication failure of the seminal Semantic blocking study (Friederici et al., 1999), where members from our lab observed an N400 even in the pure syntactic category violation condition (based on almost 200 participants) – but with an additional twist. Since our study implements true syntactic category violations and lexical-semantic anomalies, we were able to observe additive effects of the syntactic and semantic manipulations in the combined condition. Our present finding is important because *unless* semantic blocking is real, N400s elicited by combined anomalies *should* be larger than the N400s observed in response to both pure syntactic and pure semantic errors. Our results are clearly consistent with findings that individuals do not wait until syntactic-category information becomes available to process lexical-semantic information (Van den Brink & Hagoort, 2004), contrary to what is predicted by the "syntax-first" approach (Friederici et al., 2004b; Friederici & Kotz, 2003). Finally, the priming manipulation used to introduce lexical-semantic anomalies required different word-stems within target pairs. Our syntactic category violation condition therefore also reflects lexical-semantic differences: one may argue that there are no *pure* syntactic category violations, since we compare between different word-stems. One possible interpretation of the N400 in response to pure syntactic category violations would therefore be that it mostly reflects lexical-semantic activity. If so, our data would provide yet another piece of evidence against semantic blocking. This also highlights a possible limitation of our design: our data do not allow us to tease apart "pure" syntactic category violations and lexical (word-stem) effects. We will keep this limitation in mind when interpreting the ERP responses to the syntactic and semantic manipulations.

4.2. A revised time-course for the processing of local structures

We observed additive N400 effects (and no LAN effect) in response to syntactic category violations and lexical-semantic anomalies. We suggest that these findings are consistent with the increasingly popular view that the LAN and N400 are two extremes on a continuum reflecting similar processes, with a topography that can be modulated depending on the input type (Molinaro et al., 2015). In agreement studies, it has been suggested that N400s index a mismatch between word forms (e.g. *was*/**were* agreement mismatches on irregular verbs which are expressed as stems, Tanner & Van Hell, 2014) while the LAN reflects mismatching affixes (Molinaro et al., 2015). Transposing this to syntactic categories, (left-lateralized) anterior negativities may be observed when the cues to syntactic category are more grammatical in nature (i.e., provided by inflectional or derivational morphemes), while word stem cues, as we have in our experiment, could lead to more central negativities. For example, Hagoort and collaborators observed an anterior negativity in response to syntactic category violations that were cued by a morphophonological marker for a past participle in Dutch (*schroef* ‘propeller’ vs. **schroeft* ‘propelled’, Hagoort et al., 2003), while an N400 was observed in Mandarin Chinese (Zhang et al., 2013), a language where syntactic categories are specified by the word-stem. In all these cases, responses to syntactic category violations and lexical-semantic anomalies are either topographically distinct (Hagoort et al., 2003; Van den Brink & Hagoort, 2004) or additive (, and the present study) suggesting that the cognitive processes underlying these two linguistic dimensions rely on distinct neural generators.

Globally we observe an N400 in response to both lexical-semantic anomalies and syntactic category violations, and a larger N400 elicited by the combined condition. First, our finding that N400s were found for all types of anomalies clearly contradicts the semantic blocking hypothesis stipulating that semantic processing disappears once a syntactic error is identified, coherent in part with Nickels (2016) and Zhang and collaborators (2013). However, Nickels found larger semantic effects in her purely lexical-semantic anomaly condition compared to her combined condition, which is not a definitive argument against blocking. Our results do show additive effects on combined conditions, indicating that lexical-semantic processing persists even in the presence of a syntactic category violation.

Assuming that additivity of N400 effects in syntactic and semantic conditions are an indicator that these processes operate in parallel, what could these processes reflect? The N400 for syntactic category violations could index the fact that the target word carries the wrong inflectional information for the expected word-category (e.g., noun vs. verb). Given that our design relies on word-stem switches this interpretation is unlikely. A second, more conceivable interpretation, is that the N400 reflects a mismatch between the expected stem and the actual target (as suggested by Molinaro and colleagues (2015), in the context of agreement errors). On the other hand, in the lexical-semantic anomaly condition, the N400 could reflect two effects. The first effect is a facilitation (i.e. an N400 amplitude reduction) in the *correct* condition, because the target is primed. The second one would additionally index semantic integration difficulties in semantically *incorrect* conditions. Given our design, it is presently impossible to completely tease apart these two effects. However, exploratory analyses revealed that cloze probability modulated N400 effects on semantically-primed sentences regardless of whether their syntax was correct or incorrect, and that this effect was found to show a benefit for *correct* rather than a cost on *incorrect* sentence processing. Nevertheless, we must remain cautious in our interpretation as to whether these processes operate independently. Given that the N400 effects observed are moderate (in the range of 1 μV per process, such that the combined N400-effect of both sub-processes does not go beyond 2.5 μV), it is still possible that there are enough resources available to process both types of error at the same time using the same neurocognitive resources, and thus the effects would only appear to be additive. Intriguingly, this interpretation is conceivable since the N400 scalp distribution is the same for all conditions. A follow-up study with more salient anomalies may help disambiguate this issue.

We observed frontal positivities in response to syntactic category violations in the absence of lexical-semantic anomalies. Similar frontal positivities before or around 600 ms have previously been observed for pure syntactic category violations using a balanced design in English (*He made the meal to enjoy* vs. *He made the *enjoy the meal / He hoped to enjoy the meal* vs. *He hoped to *meal the enjoy*, unpublished data from Karsten Steinhauer's Neurocognition of language lab: see S1_Figure). They have also been reported in response to morphosyntactic violations such as subject-verb agreement errors (Hagoort, Brown, & Groothusen, 1993). Interestingly, these frontal positivities were shown to be insensitive to semantic manipulations (Hagoort, 2003; Martín-Loeches, Nigbur, Casado, Hohlfeld, & Sommer, 2006). It is unclear what they reflect. Some have interpreted this effect as a syntax-specific early P600, reflecting difficulty integrating a constituent within the sentence context (Molinaro, Barber, et al., 2011), or as being related to ambiguity resolution and discourse level complexity (Kaan & Swaab, 2003). These interpretations do not fit well with our own data, as the syntactic category violations are unambiguous and not more difficult to integrate than combined violations, which do not elicit a frontal positivity. Alternatively, others have interpreted this effect as a more domain-general P3a (Kasparian et al., 2017) normally reflecting surprisal and reallocation of attention (Polich, 2007). Such an effect could be explained by two features pertaining to our experimental design. First, we employ a very constraining syntactic context that makes syntactic category violations very salient (as opposed to previous syntactic category violation studies, as we argue in §4.1). Second, half our sentences have a lexical-semantic anomaly while only a quarter of them contain “pure” syntactic category violations: these violations may be more surprising in the context or our experiment and thus elicit a P3a. Since we suggested that the N400 in syntactic category violations at least partly reflects costs associated with prediction errors, the presence of a P3a could also be part of an ERP complex reflecting prediction error signal. In this case, the biphasic pattern of an N400 and an early positivity could be interpreted as a late N2b-P3a complex (Bornkessel-Schlesewsky & Schlesewsky, 2019),although this interpretation remains somewhat speculative and should be tested empirically.

Our findings that syntactic category violations elicit a significant P600, in the presence or absence of lexical-semantic anomalies are apparently in line with the widely-accepted view that the P600 reflects syntactic processing and reanalysis. Further, we observe an interaction between lexical-semantic anomalies and syntactic category violations, revealing that the effects for the two types of errors were not additive. That is, the pure syntactic category violation and the combined conditions were statistically indistinguishable. This strongly suggests that the P600s we observe – for both lexical-semantic anomalies and syntactic category violations – both rely on the same neural generators, and therefore compete for the same resources. However, since the large syntactic P600 uses up these resources, no additional amplitude increase is seen in the combined condition. This implies that the P600 in the semantic condition is the same as the one in the syntactic condition, which is incompatible with an interpretation of the P600, even in the pure syntactic category violation condition, as a reflection of structural reanalysis (Friederici, 2002). It is however completely in line with both the "Monitoring hypothesis" (Meerendonk et al., 2008) and the interpretation of the P600 as task-related component reflecting well-formedness judgements (Sassenhagen et al., 2014). We used an acceptability judgement task which encouraged the participants to consider both the grammaticality of introductory and experimental sentences and the coherence between them. Finally, the P600 mirrors behavioral responses: lexical-semantic anomalies alone were judged to be worse than correct ones, but all ungrammatical sentences were rated worse than "pure" lexical-semantic ones. These effects may simply reflect that ungrammatical sentences are easier to uniformly categorize as bad, while sentences with semantic abnormalities may induce more variability in categorization.

4.3. Individual strategies and specific responses to experimental manipulations

Recent work by Tanner and colleagues suggested that individuals rarely display a biphasic negativity-P600 profile, but rather show dominance towards an N400 or a P600 response. We used correlations to explore inter-individual data and address this possibility, as well as test whether these profiles are consistent across different experimental manipulations. First, we did not observe clear ERP response-dominance patterns in our individuals, except within the lexical-semantic condition where N400 and P600 effects were moderately and negatively correlated (r = -.54). This finding is inconsistent with previous observations in agreement (Tanner et al., 2014) and syntactic category violation studies (Nickels, 2016). The main difference between our analyses and the aforementioned studies is that we minimized component overlap by focusing on distinct electrode sites and selected time-windows that were further apart than in previous analyses (300 ms vs adjacent time-windows (Tanner et al., 2014) or 50 ms apart, (Nickels, 2016) to estimate N400 and P600 effects. Our own data suggest that response dominance should be interpreted with caution. In fact, despite variability, most of our participants displayed a biphasic response to anomalous sentences involving a syntactic category violation: that is, our participants generally seem to engage both mechanisms eliciting N400 and P600 to process syntactic category violations. In the lexical-semantic condition, the N400 and P600 effects were correlated, and while 27 participants elicited either only a N400 or a biphasic response, nine participants did not elicit an N400 at all: it seems that these participants directly engage context-integration processes, perhaps without benefiting from priming in the correct condition.

The magnitude of the P600 effect is highly correlated between all of our conditions, even the lexical-semantic anomalies and the syntactic category violations, which are different manipulations. Remember that our P600 effects reflected our acceptability ratings: it is therefore not surprising that individuals that are good at categorizing sentences as (un)acceptable can perform this task across the board, and that the cognitive processes underlying this ability are at least partly reflected by the P600. However, while the N400 effects (moderately) correlate between the combined condition and each of the two other ones, they do not correlate between the lexical-semantics and the SC conditions. Considering that we observed additive processes between Syntax and Semantics on the N400 effect, it could be that different individuals simply recruit different strategies to process these two different types of anomalies. Once again, considering that the correlation coefficients did not actually differ between comparisons, one may want to evaluate this question by implementing stronger violations that would recruit more resources and provide a better opportunity to examine the question of whether or not SC and lexical-semantic processing require shared resources.

4.4. Conclusions and future directions

The present study systematically manipulated syntactic category violations and lexical-semantic anomalies, and showed that syntactic categories are not identified first and do not condition lexical-semantic integration, providing strong evidence against strictly-serial models of sentence processing. Rather than observing a LAN-P600 complex (or eLAN-P600, or only a P600) for syntactic category violations, we systematically observed an N400-P600 one. As mentioned above, this could be due to the properties of the targets used (i.e., uninflected word-stems). Our data supports proposals where local syntactic and semantic relations are processed simultaneously. Furthermore, the N400-P3a complex elicited by syntactic category violations could reflect a prediction error response: experimental paradigms specifically addressing the additivity of the N400 and exploring the effects on the P3a may shed light on this issue. Following this, syntactic and semantic information appeared to be integrated together in discourse-related mechanisms with a later onset (P600) that were maximized by the implementation of an acceptability judgement task. Inspection of the individual data showed that it is unlikely that participants relied on one single sentence-processing mechanism either sub-served by the N400 or the P600 effect. Finally, our study showed that while N400 effects elicited by participants were not correlated across all experimental manipulations, individuals who elicited a P600 tended to do so in every condition. We suggest that this component reflects participants' ability to categorize between correct and unacceptable sentences: ongoing research from the same authors investigating online proficiency effects on ERP responses in native and second language speakers will further address this issue.

Addendum to Chapter 2. On breaking up a balanced design
(…when we can and when we should not)

1. Inspection of the sub-conditions

1.1. Introduction

The advandatages of using balanced designs have been exposed in the general introduction and in Chapter 1 of the present thesis. In Chapter 2 (and in the submitted manuscript), we commit to this design when analyzing an intepreting ERP results. However, some studies suggest that this approach can also make the data less interpretable, because (i) quantitative or (ii) qualitative differences between sub-conditions could drive the overall effect observed in response to the main conditions. Regarding (i), Mehravari, Tanner, Wampler, Valentine, and Osterhout (2015) systematically manipulated morphological complexity (The sheep **should** graze$_{\text{[simple]}}$ / **were** grazing$_{\text{[complex]}}$ in the pasture) and stimulus grammaticality (i.e. **should** *grazing / **were** *graze), and observed that the amplitude of the P600 was larger in response to morphologically complex ungrammatical stimuli, compared to morphologically simple ungrammatical ones. Qualitative differences (ii) were observed in a study by Nieuwland, Martin, and Carreiras (2013). The authors manipulated case (*Los delincuentes asaltaron* ***al***$_{\text{[marked]}}$ *chófer /* ***el***$_{\text{[unmarked]}}$ *vehículo por sorpresa*, 'The thugs assaulted the driver / vehicle by surprise') and animacy (i.e. ***al*** **vehículo*$_{\text{[inanimate]}}$ */* ***el*** **chófer*$_{\text{[animate]}}$) using a balanced design, and observed a biphasic N400-P600 response to incorrectly case-marked items. Nieuwland et al. then split between animacy sub-conditions: ***al***$_{\text{[marked]}}$ *chófer* vs. ***el***$_{\text{[unmarked]}}$ **chófer* and ***el***$_{\text{[unmarked]}}$ *vehículo* vs. ***al***$_{\text{[marked]}}$ **vehículo*. Results showed that incorrectly case-marked *animate* objects elicited N400 effects, while incorrectly case marked *inanimate* objects elicited a P600.

In the Mehravari et al.'s study, ERPs elicited by all four sub-conditions are nicely aligned until the P600 effect, so the conclusion that ungrammaticality increases the P600 effect in morphologically complex verb constructions is straighforward. In the Nieuwland et al.'s study, the ERP plots indicate sustained differences between correctly marked and unmarked conditions that start as early as the P200 component. It is therefore impossible to know whether the observed differences are driven by the errors or the correct sentences.

In the present study, as we contemplated the possibility of splitting between noun and verb sub-conditions in order to reveal potential SC-specific effect, we decided to first investigate whether there were any differences between the correct conditions.

1.2. ERP data visualisation of the sub-conditions

A subset of the data presented in Chapter 2, containing only correct sentences and syntactic category violations (Table VIII), was selected. Note that since there were 40 items per sub-condition, a good signal-to-noise ratio is preserved despite dividing up sub-conditions (Luck, 2005).

Table VIII. Sample stimuli with Condition (Correct, syntactic category violation) and Target (Verb, Noun) as factors

Condition	Target	Experimental sentence
Correct	Verb	[…] *Elles osent le <u>plaquer</u> sur le côté.*
		[…] They_{FEM} dare him <u>tackle</u> on the side.
Correct	Noun	[…] *Elles ôtent le <u>crapaud</u> sur le côté.*
		[…] They_{FEM} remove the_{MASC} $\underline{\text{toad}_{\text{MASC}}}$ on the side.
syntactic	Verb	[…] *Elles osent le *<u>crapaud</u> sur le côté.*

category violation		
syntactic category violation	Noun	[…] *Elles ôtent le *<u>plaquer</u> sur le côté.*

ERPs were plotted at representative electrodes F3, Cz, and Pz (Figure 9). Although it *seems* like syntactic category violations elicited differential effects for verb and noun targets, comparing the two correct conditions shows very clearly that the two are different almost from the beginning. Thus, differences found in the ERPs for incorrect conditions could be the result of the subtraction process of the correct from incorrect conditions within syntactic categories.

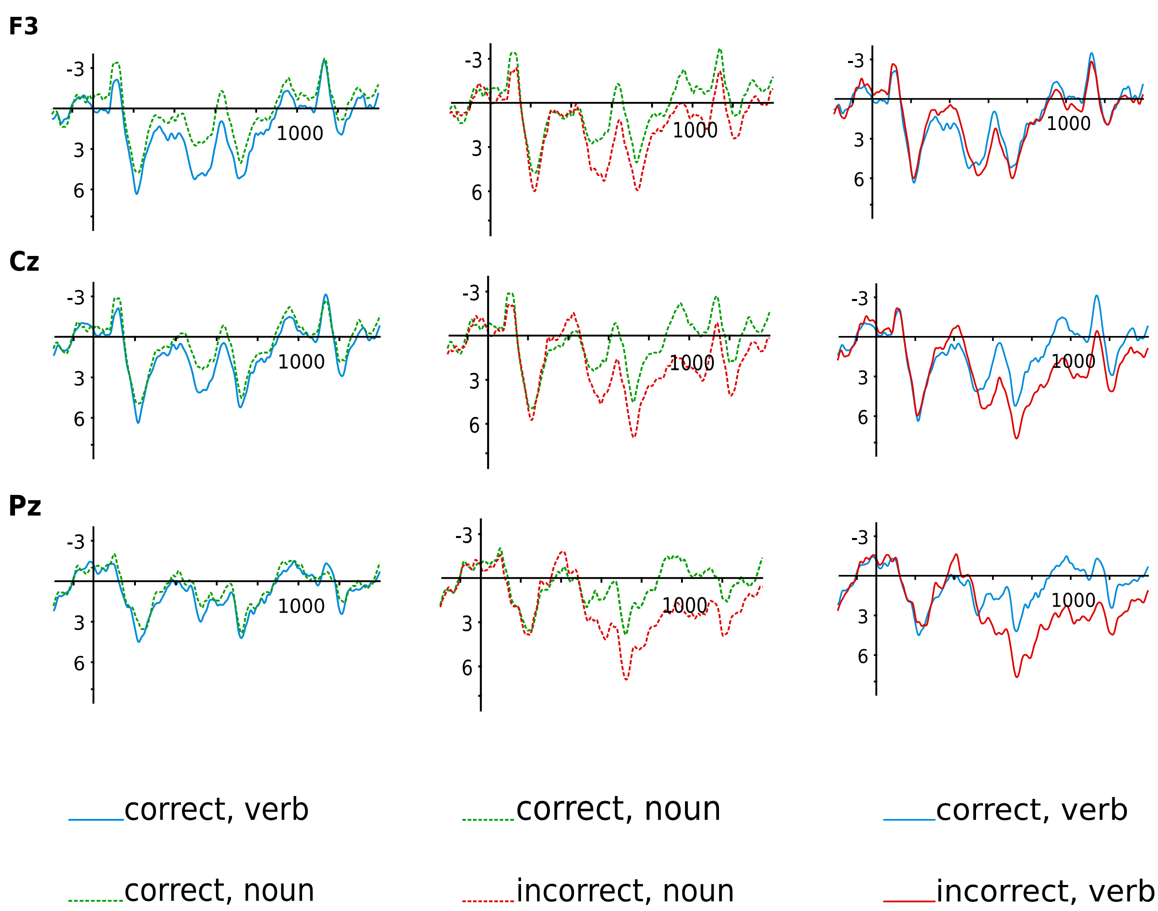


Figure 9. Average ERPs elicited at three representative electrodes (F3, Cz, Pz) in response to correct and syntactic category violation sub-conditions. *Left*: noun vs. verb correct sub-conditions; *Middle*: incorrect vs. correct noun; *Right*: incorrect vs correct verb. Target onset is indicated by the vertical bar, where tick bars represent 3 μV of activity; ticks marks on the horizontal line represent 200 ms of time.

1.3. Summary

Visualisation of the ERP shows that the correct sub-conditions are clearly different from one another. Any possible effect of syntactic category violation on either target sub-type (noun or verb) could be driven by differences in the correct conditions, by the syntactic category violation itelf, or both. However, with differences between correct conditions for each category, it is impossible to know which exactly drives possible differential effects. Therefore, in the present work, we prefer the analyses that can be performed using a balanced design, without dividing them into sub-conditions.

Visual inspection of the correct sub-conditions in Figure 9 reveals very early differences that may be due to differential processes elicited by the clitic pronoun vs. the definite determiner. We will explore this effect in the following section.

2. Retrieval and surprisal effects in the verb sub-condition

2.1. Introduction

It is not possible to break down the experimental design to examine the effect of SVCs across sub-conditions. It is, however, possible to compare between the correct sub-conditions at clitic / determiner onset, since the main verbs that preceded them matched in terms of frequency, and were phonologically similar. This would provide us information on how the two different structures are processed. The verb structure is more complex, because it involves affixation of the clitic object before the target verb and requires antecedent retrieval: evidence suggests that speakers reactivate the lemma of the antecedent in the lexicon when encourtering a pronoun (e.g., Lago et al, 2017). In contrast, although the use of a definite article should also be justified by a referent in principle, in French such articles can be used without an explicit antecedent (e.g., Royle, Fromont, & Drury, 2018).

2.2. ERP effects

Figure 10 illustrates the ERPs in the two correct sub-conditions. Target presentation is at 0 ms, and clitic / determiner presentation is at -500 ms. The baseline was therefore set between -700 and -500 ms before target presentation – that is, -200 to 0 ms after clitic / determiner.Visual inspection of the ERPs suggests that verb structures elicit a negativity that starts at frontal sites about 200 ms after clitic onset, and then is broadly distributed around 400 ms after onset. Verb targets seem to elicit a P3a starting around 450 ms after presentation. We ran mixed-effect models on two different time-windows: 200-500 ms after clitic / determiner presentation, and 450-650 ms after arget presentation. Separate models were built for midline and lateral sites. Random effects included random slopes for TARGET per PARTICIPANT, and fixed effects involved involving factors TARGET (2 levels: Noun, Verb), ANTERIORITY (Anterior, Posterior), and HEMISPHERE (Left, Right). Results are reported in Table IX.

Results for the 200-500 ms time-window relative to clitic / determiner onset showed a main effect of TARGET at both midline and lateral electrodes, indicating a broadly distributed negativity in response to clitics. Verb targets also elicited a P3a starting around 400 ms after presentation. This was confirmed by the TARGET×ANTERIORITY interaction and by follow-up models that split across levels of ANTERIORITY, which showed a main effect of TARGET at anterior electrodes (Midline: $X^2(1) = 5.09$, $p = .02$; Lateral sites: $X^2(1) = 12.57$, $p < .001$) but not as Posterior ones. Models that split across levels of HEMISPHERE showed an effect of TARGET at left sites ($X^2(1) = 18.49$, $p < .001$) only.

As a preliminary step to investigate whether these two effects reflected distinct processing strategies, we correlated the magnitudes of these two effects (negativity and positivity) across participants and observed a moderate correlation ($r = -.44$, $p = .007$). This suggest that perhaps this two mechanisms are complementary.

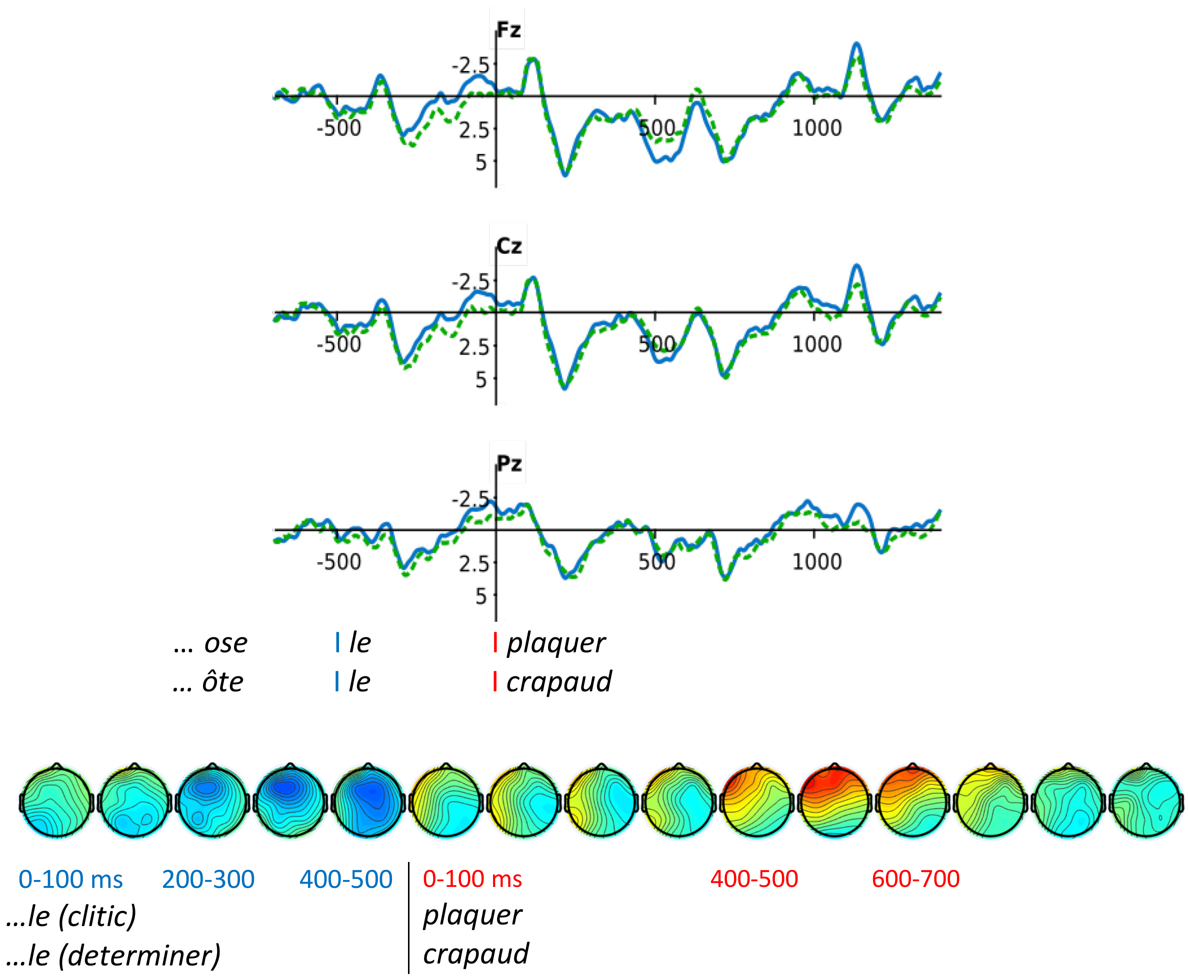


Figure 10. ERP effects of correct control verbs vs. correct transitive verbs. *Top*: Average ERPs elicited at three representative electrodes (Fz, Cz, Pz) in response to the correct sub-conditions (*blue solid line*: verbs vs. *green dotted line*: nouns). Target onset is indicated by the vertical bar, where tick bars represent 2.5 µV of activity; ticks marks on the horizontal line represent 500 ms of time. *Bottom*: Voltage maps were calculated from the correct verb minus the noun sub-conditions. Averages over 100 ms time-windows were calculated from -500 to 1000 ms relative to target onset.

Table IX. Analysis of deviance table (Type III Wald chi-square tests). Significant effects corresponding to the main mixed-effect models on average amplitudes measured in the -300-0 ms and 450-650 ms time-windows, at midline electrodes and lateral sites.

Time-window	Site	Fixed effects	Chi-square	Df	p-value
-300-0 ms *200-500 ms relative to clitic / determiner onset* (Negativity)	Midline	(Intercept)	34.112	1	< .001
		TARGET	9.134	1	.003
	Lateral sites	(Intercept)	18.706	1	< .001
		TARGET	4.278	1	.039
450-650 ms (P3a)	Midline	(Intercept)	29.189	1	< .001
		TARGET	6.458	1	.011
		TARGET ×ANTERIORITY	6.153	1	.013
	Lateral sites	(Intercept)	13.561	1	< .001
		TARGET	16.511	1	< .001
		TARGET×ANTERIORITY	10.792	1	.001
		TARGET×HEMISPHERE	7.724	1	.006

2.3. Summary

Two effects were apparent:

1- A negativity in response to clitic pronouns compared to determiners starts at frontal sites about 200 ms after presentation, and becomes broadly distributed around 400 ms;

2- A P3a is elcited by target verbs compared to nouns starting around 450 ms after presentation.

Further statistical anlayses using smaller time-windows will allow us to determine whether the frontal negativity and the more broadly disrtibuted one (which came out in the present analyses) are distinct. For now we can tentatively interpret it as lexical reactivation of the antecedent, perhaps with involvement of working memory mechanisms, as suggested by the frontal distribution around 200 ms reminiscent of an NRef (Van Berkum, Brown, Hagoort,

& Zwitserlood, 2003). The negativity and the following positivity were negatively correlated across participants. This suggests that successful mapping of "le" as a clitic pronoun engages reactivation of the antecedent, as suggested by the reviewer. Participants who engaged less in this process elicited a P3a that can be interpreted as a reallocation of attention associated with surprisal. Further analyses could investigate item effects[9].

[9] We ran some statistical analyses by splitting between 'strict' control verbs and control verbs that were most completed by nouns in our behavioral evaluations, but found no effects.

Transition. Expanding the experimental design to explore individual differences

The results presented in Chapter 2 suggest that in native speakers, syntactic category identification takes place in parallel with lexical-semantic processing at first (indexed by additive effects on the N400). Later, participants seem to engage controlled processes that include sentence reanalysis, conflict monitoring, and categorization (reflected by the P600 effect). Inter-individual variability was observed in the ERP data. Correlational analyses showed that the magnitude of the P600 effect correlated across experimental manipulations, suggesting that the same participants were engaging controlled mechanisms in response to unacceptable sentences.

In the following Chapter, we further investigate the factors that may drive inter-individual variability in native speakers and second language learners. This will allow us to test the predictions made by the convergence hypothesis, that highly proficient late second language learners may adopt native-like markers of sentence processing.

Chapter 3. Growing *Random Forests* reveals that exposure and proficiency best account for individual variability in L2 (and L1) brain potentials for syntax and semantics

Lauren A. Fromont [a,b,*], Phaedra Royle [a,b], Karsten Steinhauer [b,c]

[a] École d'orthophonie et d'audiologie Université de Montréal, Montréal, Canada

[b] Centre for Research on Brain, Language and Music, Montreal, Canada

[c] School of Communication Sciences and Disorders, McGill University, Montreal, Canada

This Manuscript has been submitted to Brain and Language and is currently under revision.

Abstract

Late second language (L2) learners report difficulties in specific linguistic areas such as syntactic processing, presumably because brain plasticity declines with age (critical period hypothesis). While there is also evidence that L2 learners can achieve native-like online-processing with sufficient proficiency (convergence hypothesis), considering all these different factors together has proven challenging. We recorded EEG while native (n = 36) and L2-speakers of French (n = 40) read sentences that were either well-formed or contained a syntactic-category error or a lexical-semantic anomaly. Consistent with the critical period hypothesis, group differences revealed that while native speakers elicited a biphasic N400-P600 in response to ungrammatical sentences, L2 learners as a group only elicit an N400. However, individual data modeling using a Random Forests approach revealed that proficiency is the one most reliable predictor in explaining the ERP responses, with N400 and P600 effects getting larger as exposure to French as well as proficiency increased, as predicted by the convergence hypothesis.

1. Introduction

While native-like first-language (L1) attainment is considered to be the gold standard for second-language (L2) learners, there is considerable variability in how individuals process their L2. They may acquire their second language early in life or during adulthood, receive various amounts of exposure, and their proficiency in different linguistic domains may vary. All of these factors play a role in cognitive processes underlying sentence comprehension in a second language, but there is an ongoing debate on their relative importance (Birdsong, 2018; Steinhauer, White, & Drury, 2009; Tanner, McLaughlin, Herschensohn, & Osterhout, 2013). Using event-related potentials (ERPs), the present study investigates online processing of French syntactic categories in L1-speakers and late (post-puberty) L2-learners of French. It addresses the relative strength of different predictors, including age of acquisition, proficiency, and exposure, in accounting for electrophysiological patterns.

Four ERP components are of interest. The ELAN (early left-anterior negativity) is an early negative shift over frontal electrodes found in response to syntactic-category violations and is thought to reflect initial phrase-structure building (Friederici, 2002, 2011). The N400 is a central negativity that is generally observed in response to difficulties in lexical-semantic retrieval (Lau, Phillips, & Poeppel, 2008), integration (Steinhauer, Royle, Drury, & Fromont, 2017), or difficulties or failure to predict target words (Kutas & Federmeier, 2011). The LAN (a left-lateralized, often anterior or sometimes temporal, negativity) and the P600 (a parietal positive shift) are observed in response to (morpho-)syntactic anomalies, such as agreement errors (see Molinaro, Barber & Carreiras, 2011, for a review, but see Tanner, 2015) or syntactic category violations (e.g., Bowden, Steinhauer, Sanz, Ullman, 2013). Although the P600 is far from reflecting *only* syntactic processes

(Bornkessel-Schlesewsky & Schlesewsky, 2008; Brouwer, Fitz, & Hoeks, 2012; Kuperberg, 2007; Meerendonk, Kolk, Vissers, & Chwilla, 2008) in L2 research, ELAN/LAN-P600 effects are often thought to be a hallmark of native-like proficiency in morpho-syntactic processing (Rossi, Gugler, Friederici, & Hahne, 2005; Pakulak & Neville, 2011). However, factors promoting native-likeness in L2 remain controversial.

Factors contributing to native-like attainment

There is a broad consensus that age of acquisition (AoA) is important when it comes to L2 learning. Later-life language learning is often associated with lower attainment attributed – at least partly – to biological factors such as decline in brain plasticity. One famous yet controversial framework, the critical period hypothesis (Johnson & Newport, 1989; Lenneberg, 1967; Newport et al., 2001), states that only during an early "critical" period in life is the brain able to establish optimal neural connections to process this input. This would mean that early L2 experience is essential to reaching native-like mastery (Birdsong, 2006; VanHove, 2013), although not all linguistic domains may be equally affected by AoA. Lexical-semantic processing seems relatively unaffected (Clahsen & Felser, 2006), even though N400 effects may be somewhat delayed or smaller in late L2-learners (e.g. Sanders & Neville, 2003; see also Mueller's 2005 review). Early ERP work on syntactic-category processing seemed to support the critical period hypothesis. Weber-Fox and Neville (1996) investigated brain responses to syntactic category violations (Max's *of proof the theorem vs. Max's proof of the theorem) and semantic violations (Max's *event of the theorem), in L1-Mandarin Chinese L2-speakers of English who were divided into five groups according to AoA. Native-like semantic N400 effects were reliably observed across groups, whereas ERP profiles for syntax violations appeared less native-like with increasing AoA. Participants who learned English after 16 did not even display a P600 component. These findings seem to suggest that late learners rely on different mechanisms than do native speakers to process syntactic structures, but show no qualitative differences in lexical-semantic processing. Similarly, auditory ERP studies that combined lexical-semantic anomalies and syntactic category violations on the same target-word (Hahne & Friederici, 2001) reported that L2-learners displayed late semantic N400s in response to these combined anomalies, while L1-speakers displayed a "syntactic" ELAN but no N400, because syntactic violations should block any lexical-semantic processing in native speakers (the "semantic blocking" effect, Friederici, Steinhauer, & Frisch, 1999). Thus, late (adult) L2-learners appear to rely more on lexical-semantic cues than syntax-guided native speakers. However, in many studies that have investigated AoA effects on syntactic processing (e.g. Weber-Fox & Neville, 1996; Hahne, 2001), AoA and L2 proficiency were correlated, which presents a challenging confound: one can question whether ERP group differences originally attributed to AoA were rather due to different proficiency levels (Caffarra, Molinaro, Davidson, & Carreiras, 2015; Steinhauer et al., 2009; Steinhauer, 2014).

The convergence hypothesis (Green, 2003**;** Steinhauer et al., 2009) proposes that different levels of language proficiency in morpho-syntax are characterized by qualitatively distinct neurocognitive processing mechanisms reflected by distinct ERP patterns that can ultimately converge on a native-like profile. It predicts that low proficiency L2-learners show no ERP responses if they do not process the error, or elicit an N400 in response to morpho-syntactic violations, reflecting the fact that they process them as lexical anomalies (see also Osterhout, McLaughlin, Pitkänen, Frenck-Mestre, & Molinaro, 2006; and Ullman's declarative/procedural model, 2001; 2004). As proficiency increases and linguistic rules become more grammaticalized, ERP signatures will transform to broadly distributed P600-like effects, and finally converge on a native-like response (e.g., a biphasic LAN-P600 effect, if that is what is observed in L1-speakers). In a study focusing on L2-learners of an artificial language, Friederici, Steinhauer, and Pfeifer (2002) showed that participants who received grammar training elicited ERP responses similar to those usually found in native speakers (early negativities followed by a P600) in response to syntactic category violations, while participants who only received vocabulary training did not (see also Morgan-Short, Steinhauer, Sanz, & Ullman, 2012). Steinhauer and colleagues (2009) also report a study where both native English speakers and high proficiency late learners elicit a LAN-P600 complex in response to syntactic category violations, while lower proficiency participants only elicit a small P600. Proficiency also seems to play a role in native monolinguals. Pakulak and Neville (2010) suggest that L1-speakers' proficiency correlates with the magnitude of both anterior negativities and late positivity in response to syntactic errors (see White, Genesee, White, King, & Steinhauer, 2006, for similar findings). In a number-agreement processing study focusing on inter-individual variability in German L2-learners, Tanner et al. (2013) observed that individuals varied along a continuum from an N400 to a P600 profile, and that only the P600 increased as participants became better able to detect anomalies.

The amount of exposure to a target language in a naturalistic environment has a positive influence on the processing of grammatical structures (e.g., Frenck-Mestre, 2002; 2001). Although few ERP studies have investigated the specific effects of exposure on syntactic processing in cohorts of L2 adult speakers, some studies included immersion in terms of years of exposure to a target language or length of residence in a country where the target language is spoken (Caffarra, Molinaro, Davidson, & Carreiras, 2015). For example, in an agreement study, Tanner, Inoue, and Osterhout (2014) found that proficiency – based on self-ratings and pencil-and-paper assessments – predicted overall response magnitude, reflected non-specifically by a larger N400 or P600 component, while immersion (length of residence) and motivation to speak like a native speaker were associated with a P600-dominant response. Note, however, that immersion in years (or length of residence) does not necessarily reflect the amount of exposure L2 speakers receive in their daily lives. This is especially true for multilingual environments, where speakers may receive some input from their target language but not necessarily use it in their daily interactions.

As individual differences along an N400-P600 continuum are also observed in L1-speakers (Osterhout, 1997; Osterhout, McLaughlin, Kim, Greewald & Inoue, 2004; Tanner and Van Hell, 2014), response dominance may be a characteristic of sentence processing in general, reflecting individual preferences to rely on one processing stream over the other, e.g., memory-based heuristics (reflected by an N400 dominance) versus procedural or combinatorial information (indexed by a P600 dominance). In sum, a biphasic LAN-P600 should not be considered the gold standard for L1 attainment. In fact, the convergence hypothesis does not predict that highly proficient L2-speakers should specifically elicit a LAN-P600 profile, nor that this profile is consistently elicited in native speakers (Steinhauer, 2009, 2014). Rather, it predicts that at very high proficiency, L2-speakers' responses to syntactic violations would be indistinguishable from those of native speakers. We believe that this proficiency-based approach would benefit from considering inter-individual variability, not only in L2-speakers, but also in L1-speakers. Tanner et al.'s (2013, 2014) studies consider L1- and L2-speakers separately, while we propose to group L1- and L2-speakers into the same analysis, without a priori assumptions.

Syntactic category identification and phrase-structure building

While the vast majority of studies observe a P600 effect in response to syntactic category violations, some studies found an additional ELAN (starting with Friederici, Pfeifer, Hahne, 1993; and Friederici, 1995), others a LAN (Van den Brink & Hagoort, 2004), and more recently an N400 (Nickels, Bokhari, & Steinhauer, 2014; Zhang et al., 2013). While the LAN-N400 discrepancy could be explained in light of inter-individual variability as explained above, we interpret with caution the much earlier ELAN effect as well as "semantic blocking", due to serious methodological concerns with experimental designs (Steinhauer & Drury, 2012). Most studies investigating syntactic category violations used "context manipulation", whereby the pre-target context differs between conditions (e.g., Hahne, 2001; Hahne & Friederici, 2001; Isel, 2007; Mueller, Hahne, Fujii, & Friederici, 2005; Pakulak & Neville, 2011; Weber-Fox & Neville, 1996, for L2 processing). For example, in French, Isel (2007) manipulated syntactic-category by omitting the noun after a definite article (*L'homme qui est dans la *Ø/maison* <u>*dort*</u> 'The man who is in the *Ø/house <u>sleeps</u>'), the context preceding the target verb <u>*dort*</u> where the ERP is analyzed. He found ELANs for syntactic category violations in both native speakers and proficient L2-learners. However, this finding is problematic because the differing pre-target contexts can elicit ERP artefacts that resemble ELANs but are not related to the syntactic category violation (Steinhauer & Drury, 2012, see also Lau, Stroud, Plesch, & Phillips, 2006**)**. In fact, Isel (2007) found an "ELAN" 100 ms after target word onset, hundreds of milliseconds *before* the syntactic violation was present in the speech signal, clearly reflecting an artefact. Nichols and Joanisse (2019) observed that in participants who learned French before puberty proficiency, not AoA, modulated the LAN response to syntactic category violations, while P600s were absent in both L1- and L2-speakers. However, they did not directly compare their conditions, so there is no statistical support for the presence or absence of ERP differences. In addition to these problems, many studies testing syntactic category violations simply fail to create syntactic or semantic violations on the target word (Steinhauer & Drury, 2012).

Using a novel paradigm in French presented in Chapter 1, we systematically manipulated syntactic-category (correct/incorrect) and lexical-semantic anomalies (primed/unprimed), while using a balanced experimental design that avoided methodological issues present in previous studies. This is one of the few studies successfully creating outright syntactic category violations *and* lexical-semantic anomalies on the target word, while deploying a balanced design where identical context *and* target words contributed equally to correct and violation conditions (see §2.2). This is a substantial improvement to context manipulation and target manipulation designs (e.g., Luo, Zhang, Feng, & Zhou, 2010; Zhang et al., 2013) and has only been implemented in a handful of studies (e.g., Bowden, Steinhauer, Sanz, & Ullman, 2013; Steinhauer et al., 2009). In Chapter 2 we reevaluated the time-course of syntactic-category identification in French native speakers (n = 36), and whether these processes interacted with lexical-semantic processing. We observed no ELAN in response to syntactic category violations, which instead elicited a biphasic N400-P600. Lexical-semantic and syntactic manipulation effects on the N400 were additive, suggesting that the two types of error were processed in parallel but independently. Using the same design, we investigated ERP responses of L2-learners of French, while comparing them to the same group of L1-French speakers, and explored effects of individual measures for AoA, exposure to French, and proficiency on ERP responses in both L1- and L2-speakers.

1.1. The current study

Previous syntactic category violation experiments yielded diverse results with respect to syntactic-category processing and its interaction with lexical-semantic processing. To better compare our data with previous findings, we will first use a traditional group design comparing L1-speakers with late L2-learners of French. While lexical-semantic anomalies should elicit similar N400s in both groups due to high levels of vocabulary mastery in the L2 group, we expect syntactic category violations to elicit a biphasic N400-P600 response at high proficiency (based on results in Chapter 2) and a reduced or even absent P600 in lower L2-proficiency learners.

Second, we will investigate the effect of nine predictors related to AoA, proficiency, language exposure, and working memory abilities on the ERP responses. In order to (a) introduce a relatively large number of correlated predictors in a single model and (b) evaluate their relative importance in explaining the modulation of ERP effects, we analyzed them using *Random Forests*, a statistical approach that has notable advantages over traditional ones but is novel to ERP research (see Tagliamonte & Baayen, 2012, and Tomaschek, Hendrix, & Baayen, 2018, for applications of Random Forests to linguistic data). The critical period hypothesis would predict that AoA is the most important predictor explaining variability in the ERP responses, with native speakers showing a biphasic N400-P600 response, and L2-learners only an N400. Based on the convergence hypothesis, we predict that proficiency will modulate the ERP profiles of L2-learners to a larger extent than AoA, and may even account for some variability within native speakers.

2. Methods

2.1. Participants

Forty native French speakers (L1-speakers) and 45 late second language learners of French (L2-speakers) who spoke Canadian or US-English as a first language participated in the experiment. Data from eight participants (four per group) were excluded from the analyses due to excessive artifacts in the EEG, and one additional participant was excluded because of technical issues during the testing session. All participants were right-handed (Oldfield, 1971, adapted to French) and reported normal or corrected-to-normal vision. Age, education, and reading habits were obtained via a short demographic questionnaire, and their working memory was assessed using forward and backward digit span tests (Soylu, 2010). Critically, we ensured that the two groups only differed on language-related measures ($p < .001$, see Table X). All L2-speakers reported an intermediate level in French: the B1 level on the Common European Framework of Reference for Languages, Council of Europe, 2011), and were not exposed to a third language more than 5% of the time. A language background questionnaire established self-reports on age of acquisition (AoA, in the case of L2-speakers this would be their first French class), age of first regular exposure to French (AoE, the moment L2-speakers started living in a French-speaking environment). The amount of exposure was estimated using two variables. Immersion (in years) is an estimation of the amount of exposure across the lifespan and was calculated by subtracting AoE from the age of the participants. Daily usage was estimated by asking the participants the extent to which they spoke French (as opposed to English) in their adult life in the following situations: at work or school, at home and during social activities. Finally, participants were evaluated on their language proficiency using a C-Test (Tremblay & Garrison, 2010) and a lexical decision task

(*LexTALE,* Brysbaert, 2013). Performance on each of the three tasks during the EEG experiment was used as independent and structure-specific proficiency measures, which according to Steinhauer et al. (2009) should be among the best predictors for ERP profiles. The internal consistency of individual responses to the linguistic tasks, estimated using Chronbach's alpha, suggests that the predictors were reliable (i.e. above the commonly accepted threshold of .7). Importantly, all measures related to AoA, exposure, and proficiency were correlated, as illustrated in Figure 11.

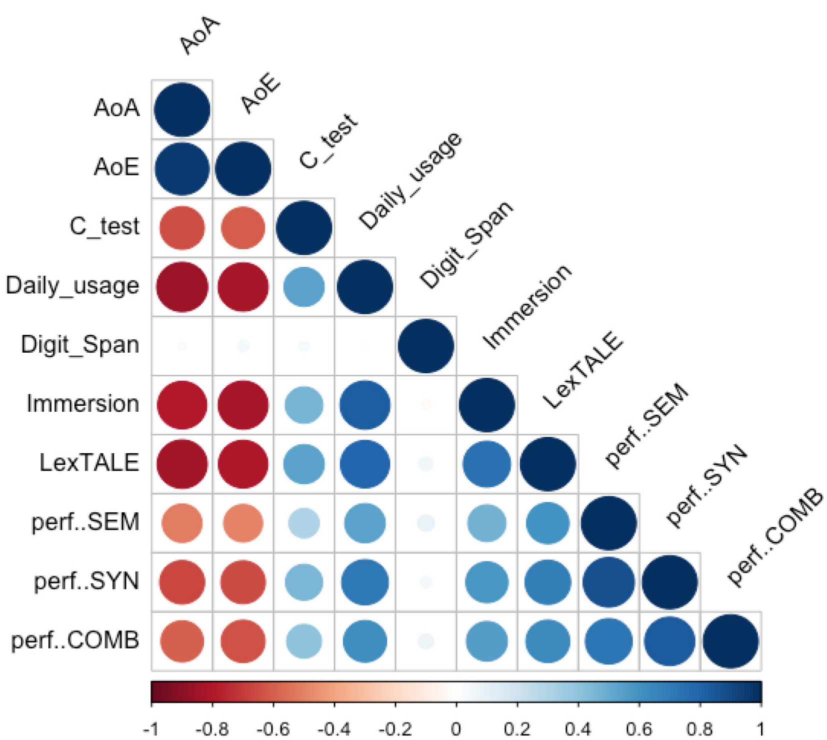


Figure 11. Correlations between the individual difference measures.
All correlations except the ones involving Digit Span were significant (*p* = .05 or lower). The magnitude of the correlation coefficient is expressed by the size and darkness of the dots; blue indicates positive correlations and red negative ones.

Table X. Participant demographics and individual measures. Significant differences are in bold.

Measure	L1 French (n = 36, 19F)		L2 French (n = 41, 26F)		t(df)	*p-value*	Internal consistency[f]
	M (*SD*)	Min – Max	*M* (*SD*)	Min – Max			
Age (years)	27.00 (5.47)	22 – 40	25.76 (4.08)	25 – 34	1.11 (76)	.270	
Education (years)	15.35 (1.92)	13 – 18	14.40 (2.86)	10 – 18	1.77 (76)	.081	
AoA (years)	**0;0 (0;0)**	**–**	**12;5 (1;11)**	**10 – 18**	**43.09 (40)**	**< .001**	
AoE (years)	**0;0 (0;0)**	**–**	**17;11 (3;6)**	**12 – 27**	**34.39 (40)**	**< .001**	
Daily usage[a]	**88;7 (10;8)**	**60 – 100**	**17;8 (13;11)**	**3 – 60**	**25.29 (76)**	**< .001**	
Immersion (years)	**27;0 (5;6)**	**22 – 40**	**7;10 (4;4)**	**2 – 21**	**16.83 (76)**	**< .001**	
Reading Habits[b]	3.88 (0.81)	1 – 4	3.98 (0.83)	3 – 5	-0.51 (76)	.609	
Handedness	81.29 (16.01)	50 – 100	82.44 (16.88)	40 – 100	-0.31 (76)	.755	
***LexTALE_FR*[c]**	**89.43 (4.04)**	**78 – 95**	**58.80 (11.03)**	**16 – 83**	**16.97 (57)**	**< .001**	.87
Cloze test[c]	**68.33 (12.47)**	**33 – 91**	**39.65 (17.56)**	**11 – 89**	**8.09 (69)**	**< .001**	.92
Performance SEM[d]	**0.28 (0.14)**	**0.03 – 0.57**	**0.12 (0.11)**	**-0.15 – 0.45**	**5.62 (76)**	**< .001**	.87
Performance SYN[d]	**0.44 (0.17)**	**0.03 – 0.78**	**0.14 (0.13)**	**-0.23 – 0.42**	**8.51 (76)**	**< .001**	.84
Performance SYNSEM[d]	**0.49 (0.18)**	**0.05 – 0.79**	**0.11 (0.25)**	**-0.52 – 0.51**	**7.84 (76)**	**< .001**	.86
Reliable Digit Span score[e]	9.94 (2.41)	3 – 13	10.00 (2.65)	5 – 16	-0.09 (70)	.926	.89

[a] Percentage of daily usage to French estimated since age 18
[b] On a scale from *0 – never reads* to *5 – reads a lot*
[c] Scores computed in percentages
[d] Scores ranging from -1 to 1 (a score of zero indicates no discrimination between correct and incorrect sentences)
[e] Task adapted from Soylu (2010). The sum of the longest string of digits recalled correctly twice, under both forward and backward conditions (Greiffenstein, Baker, & Gola, 1994).
[f] estimated using Chronbach's alpha

2.2. Materials

Table XI summarizes the conditions in our 2x2 design (SYNTAX; SEMANTICS: Correct/Anomalous). Each manipulation was based on item pairs (e.g., Condition 1 in Table XI). One item contains a verb as its target (hereafter underlined e.g., *tackle*)[10] because the preceding control[11] verb (*dare*) necessarily requires an infinitive verb as its complement. By contrast, the other item contains a noun as its target (e.g., *toad*), because the preceding transitive verb (*remove*) mandatorily selects for a noun phrase complement. Importantly, context words immediately preceding the target in both sentences were kept constant by using homographic clitic pronouns (e.g., *le* 'him') before verb targets, and definite determiners (e.g., *le* 'the') before noun targets. The implementation of homographic determiners/clitics allowed us to create syntactic category violations by swapping target words across item pairs (Condition 2), while keeping the exact same pre-target words, thus avoiding any baseline problems during ERP analysis. Experimental sentences were preceded by context sentences that licensed the use of clitic pronouns and definite determiners. The context sentences contained a prime (e.g. *hockey* primes *tackle*), allowing us to manipulate semantic context priming by swapping context sentences, such that the target was primed (Condition 1) or not (Condition 3) by its context. Note that swapping the context sentences *only* made the unprimed targets anomalous from a semantic-pragmatic point of view. We use the term "semantic anomaly" to describe this unprimed condition in order to avoid conflation with traditional priming paradigms. Finally, we created a combined anomaly condition where the

---

[10] Note that the French form of the verb 'tackle' (*plaquer*) is not homographic or homophonous with the noun 'tackle' (*plaquage*).
[11] We use the shortcuts "transitive verbs" and "control verbs" to distinguish the types of complements our main verbs take.

target word belonged to the wrong syntactic-category and was also semantically anomalous (Condition 4).

Table XI. Sample stimuli for the eight experimental sub-conditions.

Syn	Sem		Context sentence	Experimental sentence: Subject + Verb	Clitic / Determiner	Target	Prepositional phrase
✓	✓	1	*Marie et Jeanne jouent au hockey avec leur copain.* Mary and Jane are playing hockey with their friend$_{MASC}$.	*Elles osent* They$_{FEM}$ dare	*le* him	*plaquer* tackle	*sur le côté.* on the side.
		2	*Marie et Jeanne vont au marais avec leur copain.* Mary and Jane are going to the swamp with their friend$_{MASC}$.	*Elles ôtent* They$_{FEM}$ remove	*le* the$_{MASC}$	*crapaud* toad$_{MASC}$	*sur le côté.* on the side.
✗	✓	3	*Marie et Jeanne vont au marais avec leur copain.*	*Elles osent*	*le*	**crapaud*	*sur le côté.*
		4	*Marie et Jeanne jouent au hockey avec leur copain.*	*Elles ôtent*	*le*	**plaquer*	*sur le côté.*
✓	✗	5	*Marie et Jeanne vont au marais avec leur copain.*	*Elles osent*	*le*	*?plaquer*	*sur le côté.*
		6	*Marie et Jeanne jouent au hockey avec leur copain.*	*Elles ôtent*	*le*	*?crapaud*	*sur le côté.*
✗	✗	7	*Marie et Jeanne jouent au hockey avec leur copain.*	*Elles osent*	*le*	**crapaud*	*sur le côté.*
		8	*Marie et Jeanne vont au marais avec leur copain.*	*Elles ôtent*	*le*	**plaquer*	*sur le côté.*

Twenty verb pairs (transitive vs. control verbs) were selected. Our primary concern when creating stimuli sentences was to avoid pre-target and target-specific effects, as well as sentence effects that can affect the ERPs. Therefore, we ensured minimal differences between the items of each matched-sentence pair in terms of phonological and lexical properties, acceptability ratings on correct and incorrect conditions, and degree of priming between prime and target word. A total of 1280 sentences were divided into four lists using a Latin-square design, such that each participant would read 320 sentences (80 per condition), without repeating any context sentence or target word. Eighty filler sentences that were either correct or contained subject-verb agreement errors were added to each list (see Chapter 2 §2.2).

2.3. Experimental procedure

All procedures were approved by the Ethics Review Boards at McGill and University of Montreal faculties of Medicine. Participants were tested in a 2.5-hour session in the Neurocognition of Language Lab at McGill University. After completing the questionnaires, participants were seated in a chair 80 cm away from a computer monitor and read sentences in white 30-point Arial font on a black background. Each trial started with a "!!!" prompt for 1800 ms, where participants were encouraged to blink, followed by a fixation cross for 500 ms. Sentences were presented in rapid-serial-visual presentation mode: each word was presented on the screen for 300 ms followed by a 200 ms blank screen interval. At the end of every trial, 500 ms after the last word, a "???"-prompt remained on the screen until participants scored sentence acceptability by pressing a button between 1 to 5 (1: totally acceptable, 5: totally inacceptable). Their key press was immediately followed by an eye-blink prompt. After the EEG experiment, participants completed the Digit Span task, the C-test, and the lexical decision task.

2.4. EEG recording and data processing

EEG was recorded continuously from 21 Ag-Cl active-shield electrodes mounted on an EEG cap (Waveguard™ original, *ANT Neuro*, Netherlands) according to the 10-20 system (Jasper, 1958) at the following sites: FP1-FPZ-FP2-F7-F3-FZ-F4-F8-T3-C3-CZ-C4-T4-T5-P3-PZ-P4-T6-O1-OZ-O2, with a 512 Hz sampling rate and a 0.001–100 Hz online forward filter.

All EEG electrodes were referenced online against the right mastoid. An electrode between FPZ and FZ served as ground. Impedances were kept below 5 kΩ.

Data were analyzed using *EEGLab* (Delorme & Makeig, 2004) and *ERPLab* (Lopez-Calderon & Luck, 2014). Continuous data were re-referenced offline to average mastoids, and lowpass- and highpass-filtered separately with .1 and 40 Hz cut-off frequencies respectively (IIR Butterworth filters). After first epoching the data from -1000–2000ms relative to target onset, and performing a baseline correction using a 200 ms pre-target baseline interval, we rejected data that exceeded a peak-to-peak threshold of 70 µV (in 100 ms steps). We then visually inspected the remaining epochs and deleted ones still affected by artefacts. Analyses were then performed on shorter epochs of -200–1800 ms with respect to target word onset, and baseline corrected (-200–0 ms).

2.5. Statistical analyses

Behavioral and EEG data were analyzed using mixed effects models (*lme4*, Bates, Mächler, Boler, & Walker, 2015; *lmerTest*, Kuznetsova, Brockhoff, & Christensen, 2016; *emmeans*, Lenth, 2018). The maximal converging random structure included random slopes for condition per participant. Model selection was performed by decrementally removing interactions and factors from this full model until we reached the optimal model, determined by comparing two minimally different models using ANOVAs.

As performance at the task, and in particular the ability to discriminate between correct and incorrect sentences, can be viewed as the most suitable predictor variable to explore inter-individual variability for structures under investigation (Nickels & Steinhauer, 2018; Steinhauer et al., 2009; Tanner et al., 2013), we analyzed online performance as follows: From the acceptability scores (on a 5-point scale, see above) we subtracted the rating for correct from anomalous conditions, and then divided the result by 4, thereby transforming scores to an index from -1 to 1 reflecting discrimination between incorrect and correct sentences (similar to d-prime values, but for scaled data). An index of 1 means that sentences with an error were always rated 5 and correct sentences were rated 1, an index of 0 indicates that sentences were rated the same regardless of their correctness, and an (unlikely) index below 0 means that correct sentences were accepted less than incorrect ones. This index, reflecting the average performance per participant and condition was used as the input for the mixed-effects model. It was used because it is analogous to calculating difference waves, which we used to explore interindividual variability in the ERP data. In this model, we assessed the effect of CONDITION (three levels: Syntax, Semantics, Combined) and GROUP (two levels: L1 and L2).

ERP effects on the midline and lateral sites were also analyzed separately using mixed effects models. The input for the models was aggregated data[12] with one average observation per participant per sub-condition. In this omnibus analysis we employed a 2x2 design following analyses in Friederici's studies (1999 and following). We first calculated main effects and interactions for factors SYNTAX (two levels: Correct, Anomalous), SEMANTICS (two levels: Correct, Anomalous), GROUP, ANTERIORITY (two levels: F and C electrodes as Anterior sites, P and O electrodes as Posterior sites), and HEMISPHERE (two levels: Left, Right). Levels were coded such that the Intercept modeled the Correct levels of each experimental condition at the Left Anterior sites. We used ANOVA wrappers (Type III Wald chi-square test) with the *car* package (Fox & Weisberg, 2011) to present the outcomes.

[12] In Chapter 2, analyses were performed on both single observations and aggregated data and did not show any differences. We opted for aggregated data in the present study for computational resource reasons.

We also wanted to assess which predictor best explained inter-individual variability in ERP components. Although mixed-effect models could incorporate random slopes that accounted for some inter-individual variability, it was not possible to integrate our nine individual variables, because we faced multicollinearity issues whenever variables were correlated. We therefore opted for Random Forests to explore inter-individual variability. Random Forest is a machine-learning algorithm that can incorporate many variables with relatively few cases, while taking interactions into account and remaining more robust in the presence of collinearity between variables (Matsuki, Kuperman, & Van Dyke, 2016). Importantly, it allowed us to rank the importance of predictor variables in explaining the data, which is a feature particularly suited for our research questions. Psycholinguistic studies have started to successfully use Random Forests (de Aguiar, Bastiaanse & Miceli, 2016; Strobl, Malley & Tutz, 2009), but to our knowledge this is the first time they are used to analyze linguistic ERP data.

3. Results

3.1. Online performance

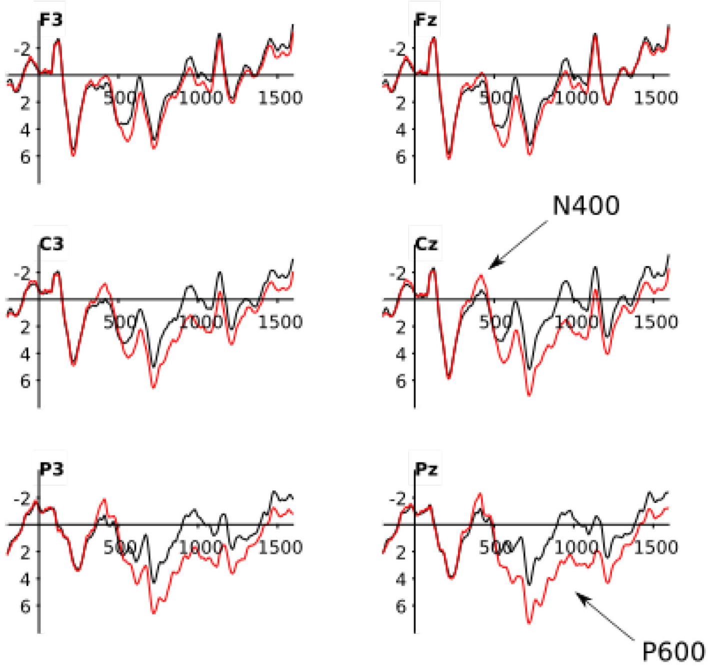
Overall, participants were accurate in judging the acceptability of sentences on the 5-point Likert scale (mean rating for correct items = 2.17, SD = 0.43; syntactic anomalies = 3.30, SD = 0.74; lexical-semantic anomalies = 2.95, SD = 0.53; combined anomalies = 3.47, SD = 0.78). Note however, that L2-speakers accepted correct sentences to a lesser extent (*Mean* = 2.35, *SD* = 0.38) that the L1-group (*Mean* = 1.93, *SD* = 0.39; *Mean difference* = 0.3; X^2 (1) = 9.88, $p = 0.002$), which impacts their performance index, described above and presented in Table X.

3.2. ERP effects: Group comparisons between L1 and L2

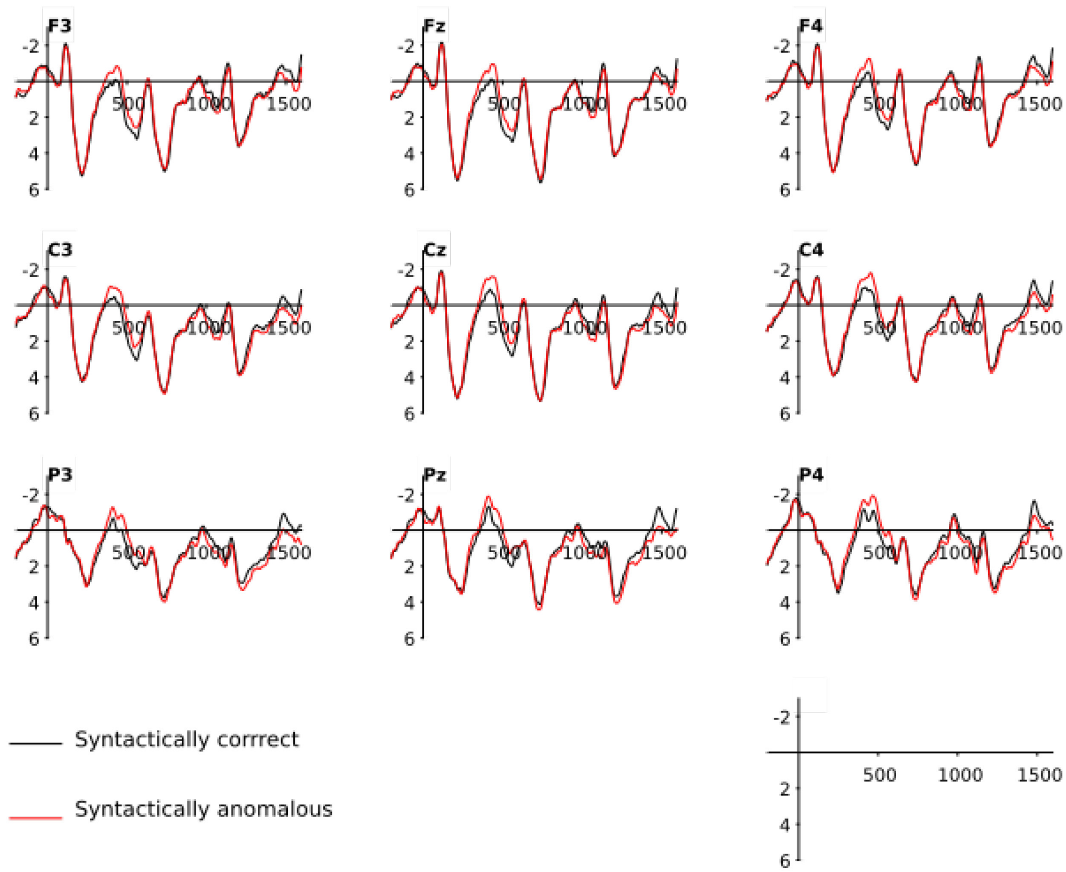
Following Friederici et al. (1999) and Nickels (2016), we first investigated the main effects of Syntax (syntactic category violations and Combined anomalies vs. Correct and Lexical-semantic anomalies) and Semantics (Lexical-semantic and Combined anomalies vs. Correct and syntactic category violations), as well as their potential interactions, comparing L1 and L2 groups (see Figure 12 and Table XII). Results showed that syntactic category violations elicited a biphasic N400-P600 response in L1-speakers, but only an N400 in L2-speakers. In contrast, semantic anomalies yielded N400s in both groups with a similar central distribution. Our statistical analyses (Table XII) tested main effects and interactions for SYNTAX and SEMANTICS in two representative time-windows (250–500 ms for N400 effects, and 800–1200 ms for P600 effects). Non-adjacent time-windows were selected in order to reduce spatiotemporal component overlap between N400 and P600 waveforms that may affect latent N400 and P600 effects and inflate correlations between them.

A- SYNTAX

L1 Group

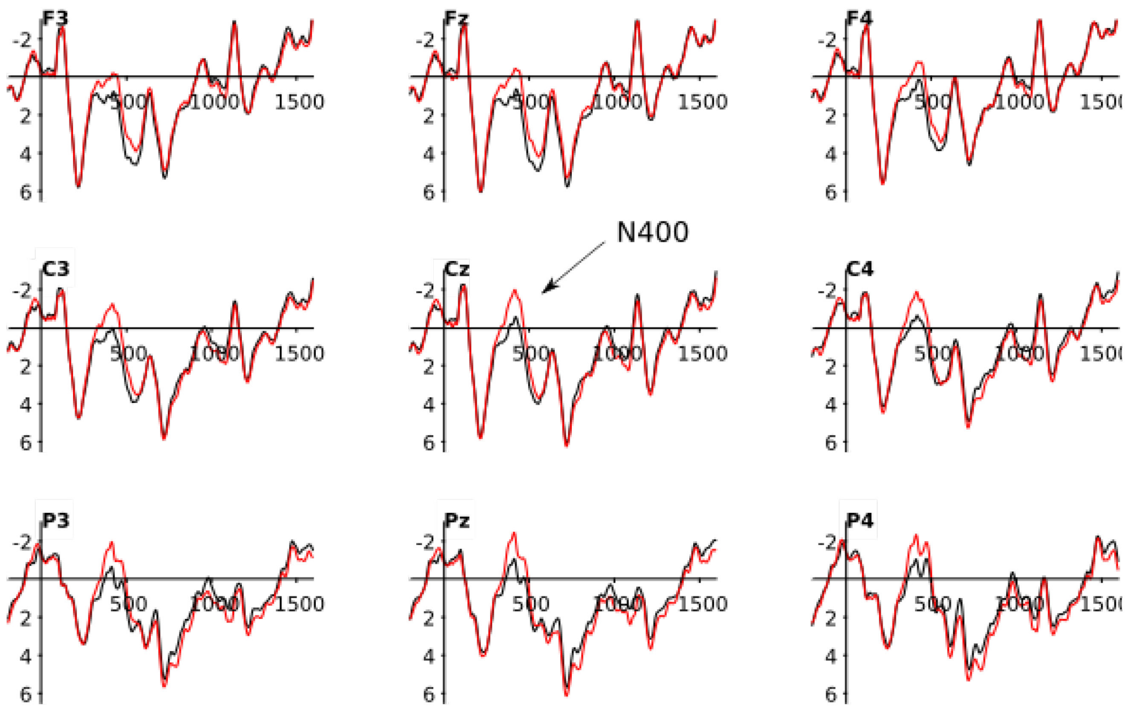


L2 Group

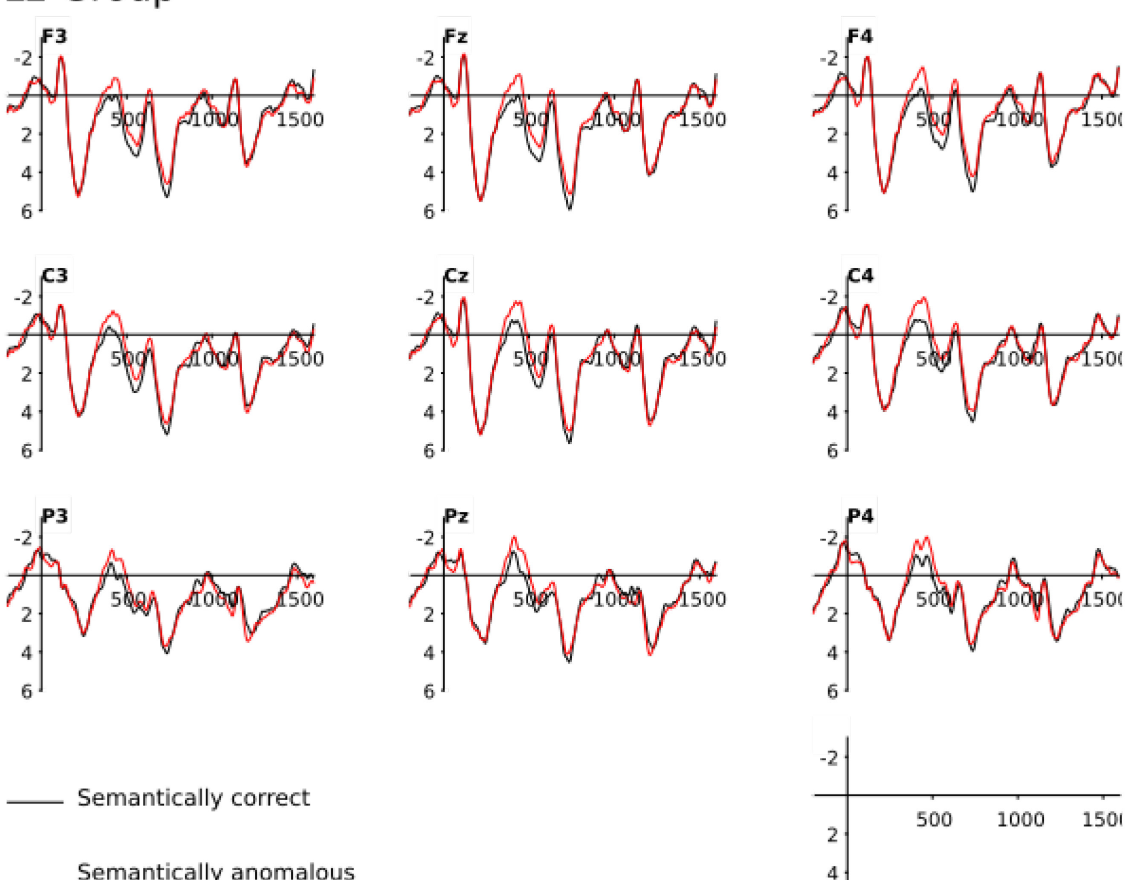


B- SEMANTICS

L1 Group



L2 Group



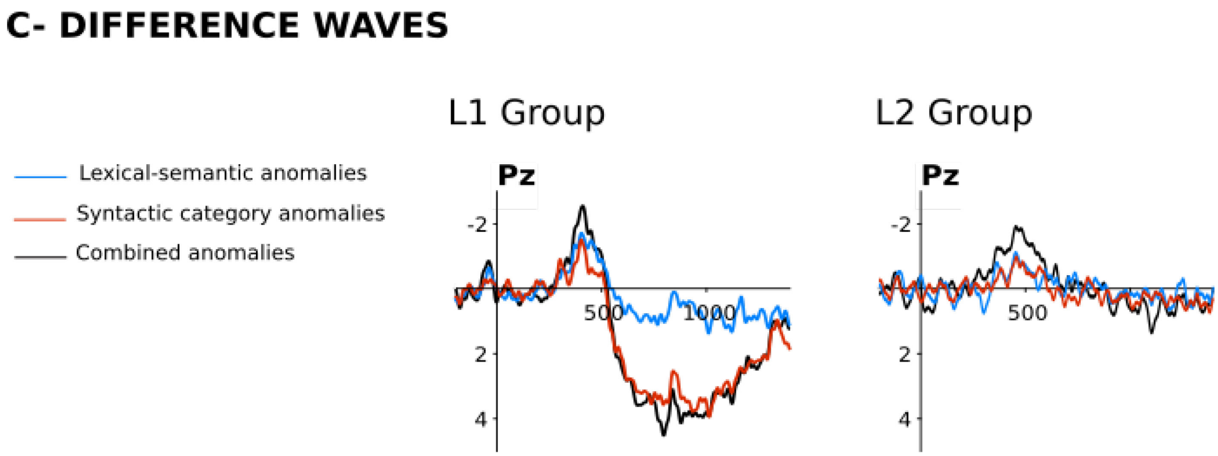


Figure 12. Grand average waveforms illustrating main effects of factors Syntax (A) and Semantics (B) in the 2x2 design in nine representative electrodes, and differences waves (C) between anomalous and correction conditions at Pz. Groups are plotted separately. Target presentation is indicated by the vertical bar, where tick bars represent 2 µV of activity; tick marks on the horizontal line represent 500ms.

Table XII. Analysis of deviance table (Type III Wald chi-square tests). reporting significant effects corresponding to the main mixed-effect models on average amplitudes measured in the N400 and P600 time-windows, at midline electrodes and lateral sites.

Time-window	Site	Fixed effects	Chi-square	Df	p-value
250-500 ms (N400 effect)	Midline $R^2 = .07$[a]	(Intercept)	25.64	1	< 0.001
		SYNTAX	6.34	1	0.012
		SEMANTICS	8.82	1	0.003
		SYNTAX×SEMANTICS×GROUP	3.94	1	0.047
	Lateral sites $R^2 = .04$	(Intercept)	32.73	1	< 0.001
		SYNTAX	3.84	1	0.050
		SEMANTICS	7.28	1	0.007
800-1200 ms (P600 effect)	Midline $R^2 = .05$	(Intercept)	0.46	1	0.493
		SYNTAX	17.22	1	< 0.001
		SEMANTICS	5.52	1	0.019
		SYNTAX×GROUP	7.61	1	0.006
		SYNTAX×ANTERIORITY	5.64	1	0.018
		SYNTAX×GROUP×ANTERIORITY	5.64	1	0.017
	Lateral sites $R^2 = .06$	(Intercept)	1.59	1	0.207
		SYNTAX	7.38	1	0.007
		SEMANTICS	6.91	1	0.009
		SYNTAX×ANTERIORITY	20.68	1	< 0.001
		SYNTAX×GROUP×ANTERIORITY	22.3103	1	< 0.001

[a] We report here the marginal R-squared values that are associated with the fixed-effects.

In the 250–500 ms time-window, both L1- and L2-speakers showed broadly distributed N400s in response to lexical-semantic anomalies (i.e., a main effect of SEMANTICS), and a similar (albeit smaller) N400 effect for syntactic violations (main effect of SYNTAX). The absence of interactions between the two factors suggests that these N400 effects are additive. Additive semantic and syntactic N400 effects would predict the largest N400 in the combined violation condition, and this is exactly what was found in both groups (see Figure 12C). The SYNTAX×SEMANTICS×GROUP interaction at midline electrodes did not reveal any significant follow-up interactions by SYNTAX, SEMANTICS or GROUP (despite a marginal SYNTAX×SEMANTICS interaction in the L2 group, $p = .06$), and should be interpreted with caution. It primarily reflects larger N400s in L1 than L2 group for pure syntactic and semantic anomalies, but comparable N400s in both groups for the combined condition (as illustrated in Fig 2C).

In the 800–1200 ms time-window, we found a small but significant positivity (at Pz: 0.55μV for L1 and 0.32μV for L2) in response to lexical-semantic anomalies in both groups (main effect of SEMANTICS), and a large P600 effect in response to sentences with syntactic errors (main effect of SYNTAX) that was more posterior and mostly carried by native speakers, as supported by significant SYNTAX×GROUP×ANTERIORITY interactions especially at lateral sites. Follow-up analyses within each group confirmed this pattern: a SYNTAX× ANTERIORITY interaction was found in L1-speakers: Midline: X^2 (1) = 9.84, $p = .002$, Lateral sites: X^2 (1) = 50.97, $p < .001$, while no corresponding effect was observed in L2-speakers. In L1-speakers, the P600 was larger at posterior (Midline: SYN–COR = 2.56μV, $t(43.14) = 8.176$, $p < .001$; Lateral: SYN–COR = 2.27μV, $t(41.6) = 10.17$, $p < . 001$) than anterior sites (Midline: SYN–COR = 1.16μV, $t(43.14) = 3.711$, $p < .001$; Lateral: SYN–COR = 0.60μV, $t(41.6) = 2.79$, $p = .008$).

As pointed out by one of the reviewers, observing a P600 for L2 learners for Semantics rather than Syntax seems rather surprising and, moreover, does not seem to be reflected in the ERP plots (Figure 12C). As we will see, this effect is related to individual differences, which we will address next.

3.3. Exploring inter-individual differences in ERP responses

3.3.1. Selecting a dependent variable

Previous work on inter-individual differences has pointed to a negative correlation between N400 and P600 effects, which motivated the computation of a response dominance index (N400 or P600 dominance) and a response magnitude index (response amplitude regardless of the component) (Tanner et al., 2014). Recall that a 'negative correlation' means that when the N400 gets smaller (i.e., more positive), the P600 gets larger (also more positive), such that this pattern could be interpreted by (i) component overlap and (ii) a single slow wave modulating both ERP components (thus pointing to fundamentally different data interpretations). In fact, some studies consider the entire 'N400-P600' complex and incorporate time as a variable using generalized additive modeling to better explore the effects of AoA on inter-individual variability (de Aguiar, Bastiaanse, & Miceli, 2016; Strobl et al., 2009). To determine the appropriateness of these approaches (and their conclusions), one would need to take topographic differences between the N400 and P600 into account. However, the aforementioned studies collapse across regions of interest to estimate the ERP patterns, running the risk that the correlation between N400 and P600 simply reflects an overlap between these components. To determine whether we should consider P600 and N400 effects separately, or rather consider response dominance and magnitude indices as our dependent variable, we first calculated how our components correlated across the three incorrect conditions. Unlike Tanner et al. or Meulman et al., we selected C3-Cz-C4 for the central N400 effect, and P3-Pz-P4 for the posterior P600, as these recorded maximal effects for each component, respectively. In addition, we minimized the impact of (trivial) component overlap by using non-adjacent time-windows (250-500ms for the N400, 800-1200ms for the

P600). A significant correlation between components was only found in the semantic manipulation condition ($r = -.46$, $p < 0.001$). Considering the two components separately is therefore a more appropriate approach to investigate our data.

3.3.2. Random forest methodology

We evaluated the strength of all our language-related variables, including online performance, in predicting ERP effects elicited by each incorrect condition. We grew six Random Forest models for each of the three difference waves (the 3 incorrect conditions minus the correct condition), in the above-mentioned time-windows for the N400 (at C3-Cz-C4) and P600 (at P3-Pz-P4). Our method was adapted from Tomaschek et al. (2018) and used unconditional variable importance with the *ranger* package (Wright & Ziegler, 2017). This method is deemed superior to both (i) the conditional variable importance implemented in the *party* package (Hothorn & Zeileis, 2015) which is heavier on resources, and (ii) the *randomForest* package (Liaw & Wiener, 2002), which inflates the importance of continuous variables and correlated data (Strobl, Boulesteix, Kneib, Augustin, & Zeileis, 2008; Strobl, Malley, & Tutz, 2009). The method implemented with *ranger* should not be biased against Group (our categorical variable), or promote highly correlated variables (e.g. performance over Digit Span score)[13]. The number of variables randomly chosen at each node (mtry), as well as trees that are grown (num.trees) can both influence the outcome of random forest models. We thus determined the optimal values for these parameters using the train( ) function of the *caret* package (Kuhn, 2008). Prediction accuracy was evaluated under 10-fold cross-validation. A model with optimal values was then calculated. In order to ensure replicability, a

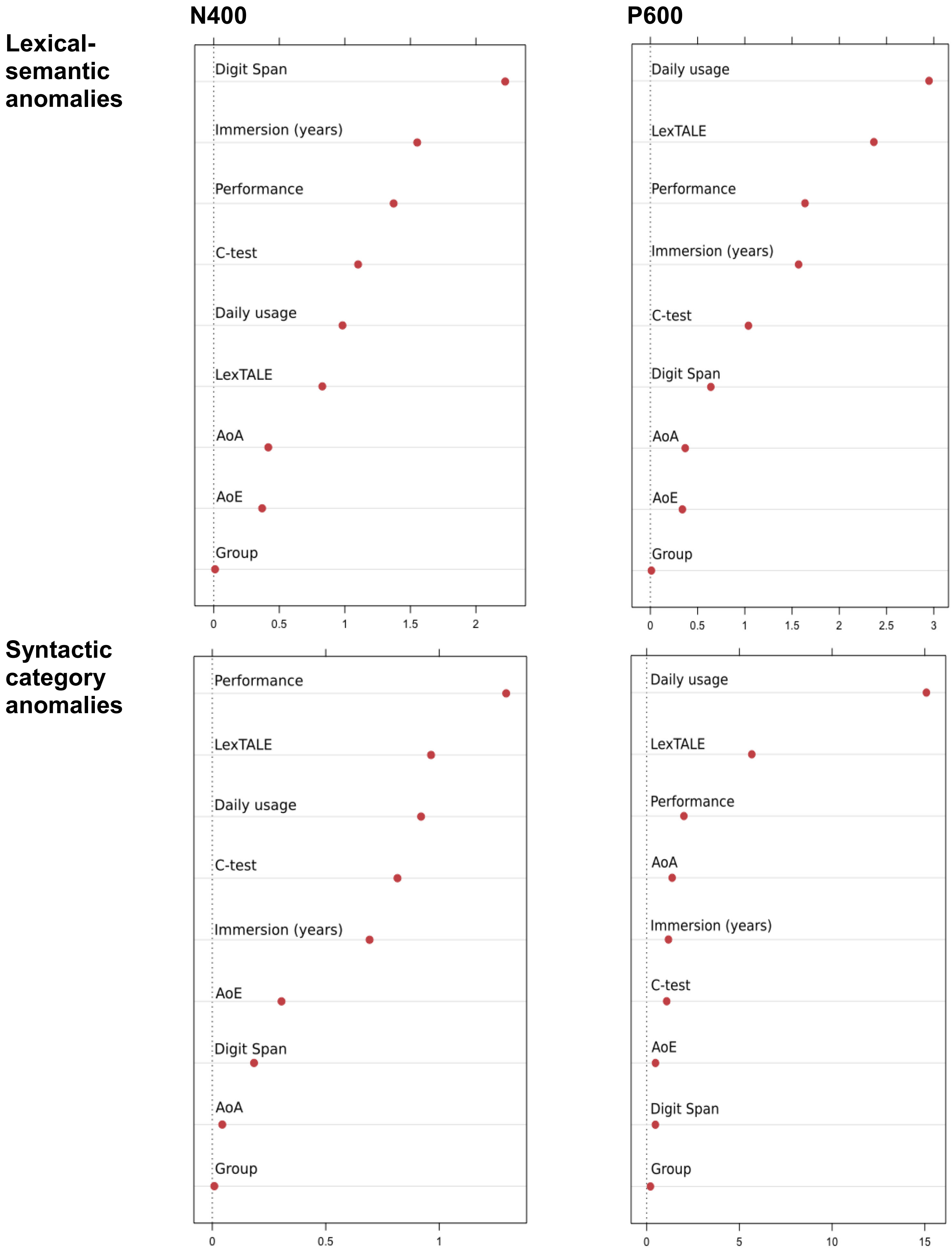
[13] We thank Darren Tanner for raising this concern with our initial analyses using the *randomForest* package.

random seed was set (at 19). To estimate variable importance, the algorithm randomly selected subsets of the data and modeled the effect of each predicting variable in every subset. Accuracy of each prediction was compared to the remaining observations. Strength of a predicting variable was calculated by randomly permuting its levels and thus erasing its importance: a predictor is deemed important if the model becomes worse (Breiman, 2001).

Following suggestions made by the reviewers, we considered the possibility that even though Random forests deal with multicollinearity better than linear regressions, the effects of predictors related to exposure and proficiency may be conflated with group effects. In order to further tease these effects apart, we (1) ran Random forests for each group and condition and (2) used conditional inference trees to illustrate how the most important variables interact. As no tree function is implemented in the *ranger* package itself, we used the ctree( ) function of the *party* package (Hothorn, Hornik, & Zeileis, 2006; Hothorn & Zeileis, 2015). Trees predict the value of continuous variables (ERP amplitudes) from a set of continuous or categorical predictor variables, using recursive binary partitioning. Trees provide estimated "split points" at which the nodes separate between two groups with different outcomes. The splitting criteria are calculated using the permutation-test framework (Hothorn et al., 2006). For each possible split, the test-statistic value is calculated under a certain label rearrangement: if they are interchangeable, the splitting value is not relevant (and would not be reported by the software). Variables included were selected using a backward elimination procedure that compared out-of-bag R-squared values. Backward elimination procedures have been adopted in gene selection research (Díaz-Uriarte & De Andres, 2006) and psycholinguistic studies (de Aguiar, Bastiaanse, & Miceli, 2016). Random forests using ranger() or cforest() are superior to individual trees using ctree() to account for the variance of the data; nevertheless, trees provide useful insight on the complex interactions that characterize the data. All the following models and their outputs are available in the supplementary materials.

3.3.3. Predictor variable importance

To make comparisons across conditions easier, we present variable importance for each ERP response and condition in Figure 13. Group, AoA, and AoE were the least important in almost all conditions, except for the P600 response to combined anomalies (in this case they were mildly important). Immersion (in years) was a very important predictor for ERPs to Semantic anomalies, and mildly important in the Combined condition. Importantly, daily usage (in percentage) strongly predicts the P600 effect in all conditions. Proficiency measures (online Performance, LexTALE, and C-test scores) were also very important predictors for all conditions. Finally, Digit Span score was an important predictor for the N400 effect in the Semantics condition.


N400
P600
Lexical-semantic anomalies
Digit Span
Immersion (years)
Performance
C-test
Daily usage
LexTALE
AoA
AoE
Group
0
0.5
1
1.5
2
Daily usage
LexTALE
Performance
Immersion (years)
C-test
Digit Span
AoA
AoE
Group
0
0.5
1
1.5
2
2.5
3
Syntactic category anomalies
Performance
LexTALE
Daily usage
C-test
Immersion (years)
AoE
Digit Span
AoA
Group
0
0.5
1
Daily usage
LexTALE
Performance
AoA
Immersion (years)
C-test
AoE
Digit Span
Group
0
5
10
15

Combined anomalies

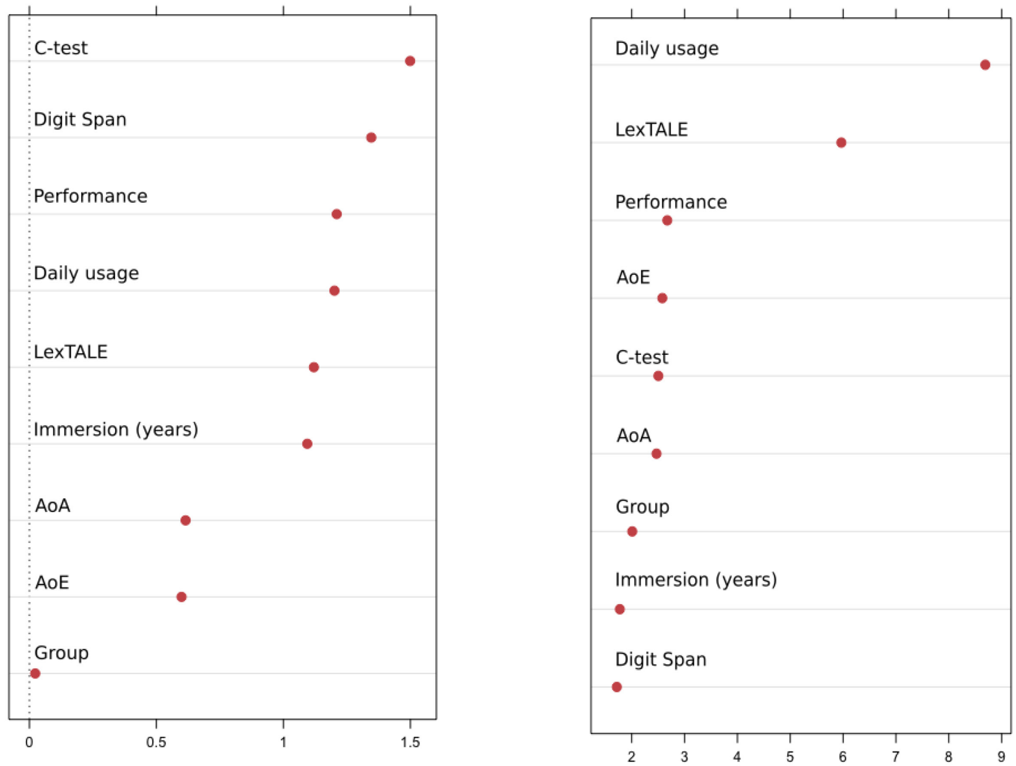


Figure 13. Predictor importance for N400 and P600 effects in all conditions.

Variable importance for the N400 and P600 effects across all conditions are illustrated for each group in Supplementary materials. The purpose of these analyses was to investigate whether variable importance differed between groups, and in particular whether AoA would be important for L2 speakers. In fact, even when focusing the analyses on the L2 group alone, AoA does not emerge as an important predictor. Not only in the L2 group, but also in L1 participants, the N400 effect was modulated by daily usage and proficiency measures (i.e., LexTALE score and online performance). In the L2 group, C-test and digit span scores were also among the most important measures. For the P600 effect, however, there is a contrast between the L1 group, where lexTALE matters most, and the L2 group, were daily usage is the most important predictor.

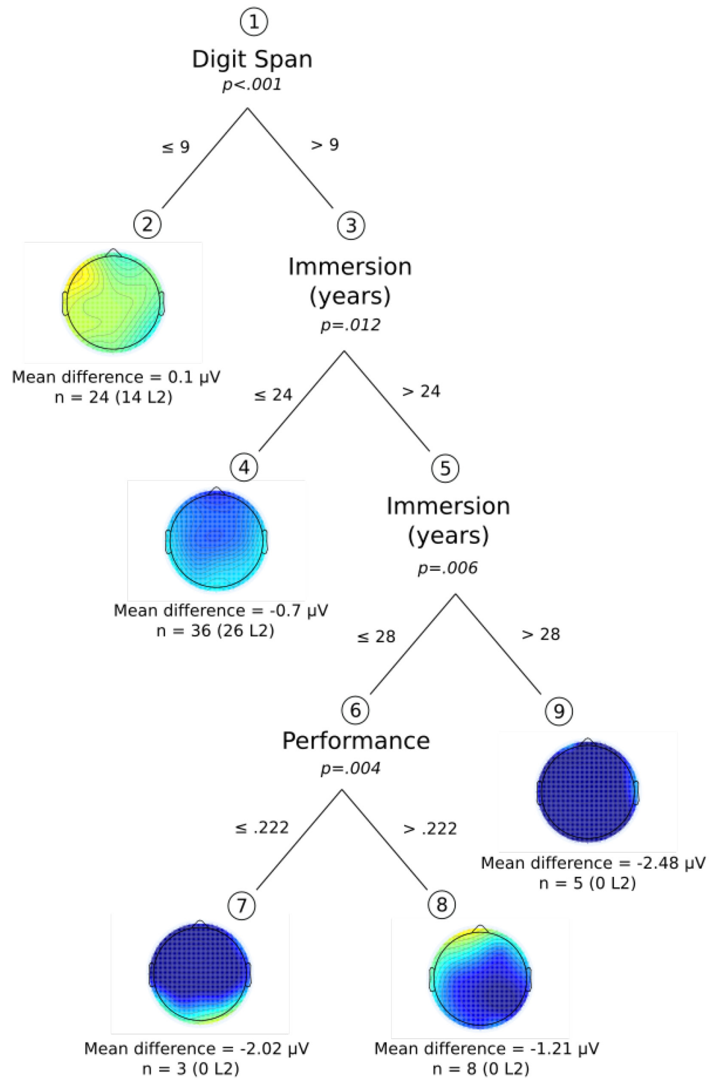
3.3.4. Illustration of the effects using decision trees

To illustrate the effects expressed by the individual trees, we drew scalp maps based on split points determined by the inference trees in Figure 14. At split points, we indicate how many speakers (and how many L2 learners) fall into each group. In the Semantics condition, the N400 effect is first determined by digit span (scores below 9 do not show any effect). Then, it is split by immersion: the lowest immersion cluster has a mix of L1 and L2 speakers, but the higher clusters (> 24, nodes ⑤ to ⑨) only have L1 speakers. Finally, performance had an effect in a higher-immersion subgroup (node ⑥), but higher performance was actually associated with smaller N400s (node ⑦ vs. ⑧). No effect was apparent for the P600. In the Syntax condition the N400 effect was best explained by performance: the split point revealed that a few L2 participants who were at chance level did not display an N400 effect (node ②), but a frontal positivity instead. The P600 effect was first split by daily usage: only L1 participants fall into the very high group (node ⑦). Among them, surprisingly, participants who had higher LexTALE scores show a *smaller* P600 (node ⑧ vs. ⑨). Participants who were exposed to French between 22% and 90% of the time also display a P600 (node ⑥). For participants with the lowest daily usage (L2 speakers exclusively) online performance determines the presence or absence of a (small) P600 — lower performance is once again associated with a frontal positivity (node ④ vs. ⑤). In the Combined condition no split was observed for the N400, while daily usage significantly accounted for the P600, with participants above 31% daily usage displaying the largest effect (node ⑤). Surprisingly, digit span, a less important variable, showed a split among participants who had lower exposure, with larger digit span reflected by a negativity (node ③ vs. ④).

Lexical-semantic anomalies

-

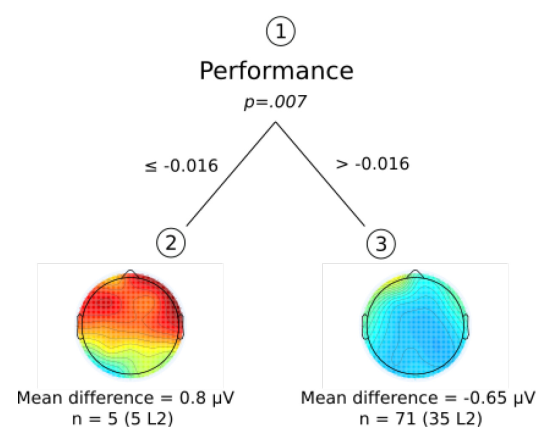
N400



Syntactic category anomalies

-

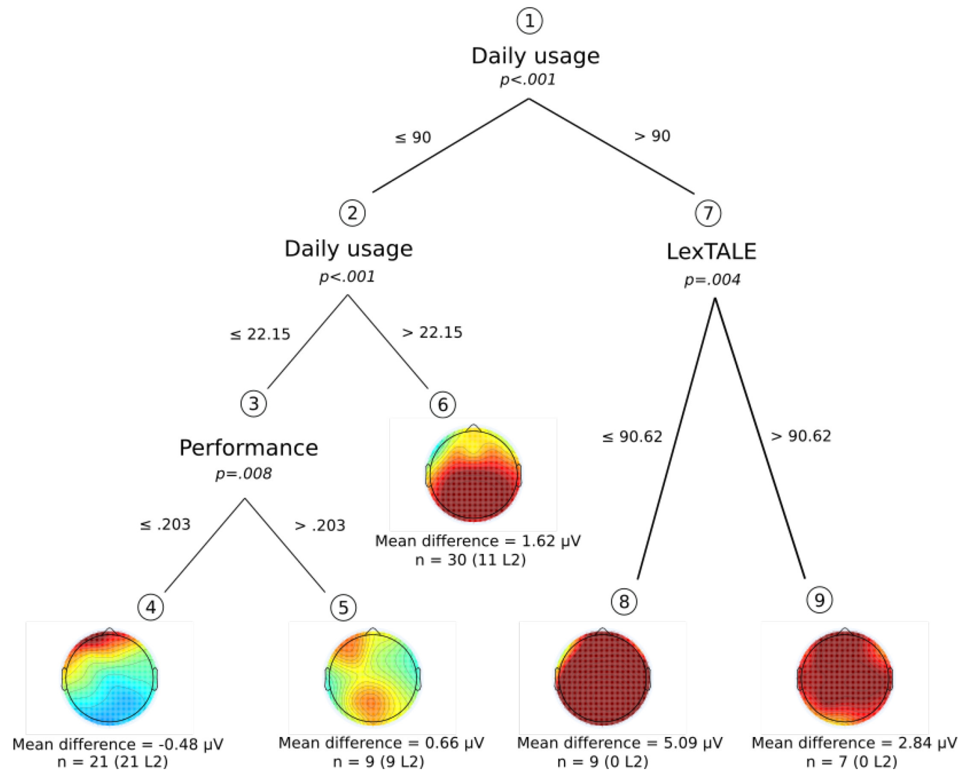
N400



Syntactic category anomalies

-

P600



Combined anomalies

-

P600

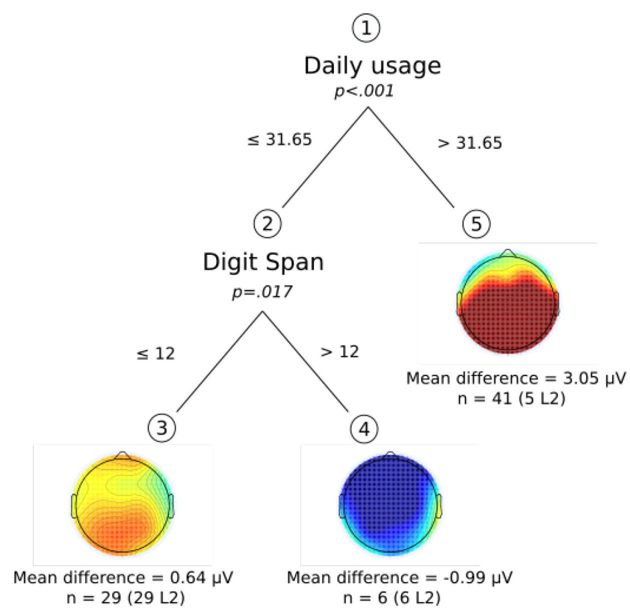


Figure 14. Maps illustrating inference tree outputs for Semantic, Syntactic, and Combined anomalies in N400 and P600 time-windows. Split points were determined using inference-tree calculations (the circled numbers are the nodes). Under each map are mean amplitude values estimated by the models, as well the number of participants who fell into each cluster (in parenthesis the number of L2-speakers).

4. Discussion

The present study used ERPs to investigate the time course of cognitive processes in response to syntactic category and semantic priming manipulations as well as their interactions, in both first and second language speakers of French. Since most previous studies investigating syntactic-category processing used unbalanced designs (Hahne, 2001; Hahne & Friederici, 2001; Isel, 2007; Mueller et al. 2005; Pakulak & Neville, 2011; Weber-Fox & Neville, 1996) that could potentially lead to artefacts (Steinhauer & Drury, 2012), our first goal was to reevaluate this issue at a group level. We thus employed a design in French that manipulated both contexts and targets while systematically controlling for them. We observed that both L1- and L2-speakers elicited similar N400s to syntactic and semantic manipulations. While only L1-speakers elicited a reliable P600 to ungrammatical sentences as a group, semantic anomalies elicited small P600s in both groups. However, this pattern was not observed when focusing on individual differences in our data: the observed pattern will be expanded below. Second, we investigated what factors related to language learning could account for the range of observed ERP profiles. We estimated the relative importance of our predictors using Random Forests, as well as decision trees to assess the significance of the most important factors identified. We found that daily usage, immersion, and performance on the online acceptability-rating task were the most reliable predictors explaining our data. Specifically, these factors predicted the amplitude of both the N400 and the P600 effects, in L1- and L2-speakers *alike*. Immersion (but also importantly digit span) predicted the semantic N400, while daily usage and performance were better predictors for the syntactic N400 and the P600 effect.

4.1. Group effects

Our finding that syntactic category violations elicit an N400 instead of an ELAN in both groups differs from a large body of literature in L2-processing (Hahne, 2001; Hahne & Friederici, 2001; Isel, 2007; Mueller et al. 2005; Pakulak & Neville, 2011) as well as L1 studies (see Friederici, 2011, for a review, and Chapter 2 of the present thesis for additional discussion of sentence-processing results). We attribute the absence of an early effect to our balanced[14] design, which avoids baseline issues. Together with Steinhauer and Drury (2012), this is evidence that the ELAN cannot be viewed as a reliable index of native-like linguistic achievement in L2-speakers of French. The absence of a LAN also differs from previous studies (Nichols & Joanisse, 2019; Steinhauer et al., 2009), although some studies focusing on L1-processing have observed an N400, and no LAN, in response to syntactic category violations (Nickels et al., 2014; Zhang et al, 2013). In morphosyntactic studies focusing on agreement, there is currently a debate on whether the LAN is, in fact, an N400 that adopts a frontal, left-lateralized topographic distribution when it is superimposed by a right-posterior P600 (Guajardo & Wicha, 2014; Tanner et al., 2014; but see discussion in Courteau et al, 2019 and Royle & Steinhauer, in prep). Since some LANs have been observed in the absence of a P600 (Hasting & Kotz, 2008), and considering recent studies suggesting that at least some LAN effects are not a product of averaging (Caffarra, Mendoza, & Davidson, 2019), one possibility is that LAN and N400 are not quite distinct but represent a continuum reflecting a

[14] One reviewer suggested that analyzing the sub-conditions separately (i.e., breaking up the balanced design) could reveal qualitative (Nieuwland, Martin, & Carreiras, 2013) or quantitative (Mehravari, Tanner, Wampler, Valentine, & Osterhout, 2015) differences that may help understand the data. As splitting between sub-conditions revealed differences between the *correct* sub-conditions (i.e., control vs. transitive verbs), it is impossible to interpret whether the ERP effects of syntactic anomalies in separate sub-conditions are driven by context or lexical effects. Extended discussion can be found in the Addendum to Chapter 2 and in Steinhauer, Drury, Portner, Walenski, & Ullman (2010).

mismatch between predicted features and the target, with a topography becoming more N400-like when these features belong to the lexical-semantic domain (Bornkessel-Schlesewsky & Schlesewsky, 2018; Molinaro, Barber, Caffarra, & Carreiras, 2015). In the context of our experiment, the N400 could indicate that syntactic-category identification in French relies more on word stem information than morphological regularities.

Even without considering word class markers, processing the word stem of a verb (e.g., *tackle*) replacing a noun (e.g., *toad*) – and *vice versa* – can be expected to result in a lexical-semantic mismatch and a corresponding N400, unless semantic processing is assumed to be "blocked" by a syntactic category violation on the same word (Friederici, 2002 and Friederici et al., 1999; 2004). Crucially, our finding that enhanced N400s for both syntactic and semantic anomalies were additive (in both L1 and L2) clearly contradicts the "semantic blocking" hypothesis stipulating that semantic processing becomes moot once a syntactic error is identified, as already predicted by Steinhauer and Drury (2012) and partly confirmed by Nickels et al. (2014) and Zhang et al. (2013). Studies that seemed to support semantic blocking (e.g., Friederici et al., 1999; Hahne & Friederici, 2002) invariably used a paradigm in which the lack of an N400 in the "combined violation" condition could also be explained without any reference to the syntactic violation (see Steinhauer & Drury 2012 for details). In contrast, our semantic manipulation using contextual priming is immune to this concern. The presence in of a clear semantic N400 whose amplitude was *enhanced* rather than reduced by a simultaneous category violation, provides very strong evidence that lexical-semantic processing is not blocked by syntactic category violations. From this perspective, the absence of N400s in certain "double-violation" conditions from Friederici's group was likely due to a consistent methodological flaw in creating that condition, and not to semantic blocking.

Moreover, since our paradigm was a reading study in which all word information became available at once and did not unfold over time as in auditory research, we can also rule out Friederici and Kotz's (2003) creative hypothesis that the word stem's status as a potentially free or bound morpheme may determine whether semantic blocking takes place or not. Instead, we conclude that semantic blocking is a myth, based on misinterpreted data. Native speakers process semantic and syntactic information in parallel, and late L2-learners converge toward the same pattern at high exposure and proficiency levels.

The finding that lexical-semantic anomalies elicit an N400 is in line with a very prolific body of literature (reviewed by Kutas & Federmeier, 2011). The observed priming effects on the N400 are likely to reflect distinct processes (namely, automatic spreading activation and post-lexical integration, see Steinhauer et al., 2017, for a recent account), but our design does not allow us to tease these apart, and we consider priming effects as a whole. We found no interactions for syntactic category errors and priming, suggesting additive effects of the lexical-semantic and syntactic manipulations. Additive ERP effects in turn suggest that the underlying cognitive mechanisms used to process these different types of information are neurally and functionally distinct. This finding coheres with literature arguing that different neurocognitive and neurolinguistic mechanisms reflected by modulations of the N400 effect can be independent. For example, Chow et al. (2014), found additive effects on the N400 when manipulating word repetition and predictability, given a sentence context (*Brian looked all over the house for his missing keys/watch before leaving for work*, where target words were expected or not (*watch*), and old versus new). We could therefore interpret the N400 in response to syntactic category violations as a mismatch between predicted features (e.g. *–er, –ir, –re* endings in verbs) and the target (as proposed by Molinaro et al., 2014; or Tanner et al., 2014), while the N400 in response to lexical-semantic anomalies could reflect the absence of priming. Note that the additive effects should, however, be interpreted with caution: as the amplitude of the N400 in every condition is relatively small, it could mean that not all the resources are recruited to process either type of anomaly, and that they therefore do not need to compete when the two anomalies are combined (see Chapter 2 for further discussion).

Consistent with the majority of the literature on L1 and L2 processing, syntactic category violations elicited a larger P600 in L1- than in L2-speakers (e.g. Hahne & Friederici, 2001; Rossi et al., 2006; Steinhauer et al., 2009; Weber-Fox and Neville, 1996). This P600 effect is compatible with virtually all previous P600 accounts, ranging from task-related well-formedness judgments (Sassenhagen et al., 2014), conflict resolution in a monitoring context (Vissers et al., 2008), and sentence diagnosis and reanalysis (e.g. Friederici, Mecklinger, Spencer, Steinhauer, & Donchin, 2001). A P600 effect was also found for pure lexical-semantic anomalies, even though the ERP plots suggest that the effect is small (therefore the degree of additivity of the P600s should be viewed with caution). This finding is in line with previous studies involving judgments (Diaz & Swaab, 2007; Royle et al, 2013; Steinhauer, Drury, Portner, Walenski, & Ullman, 2010). Using the same materials but focusing only on the L1 participants, we had observed a weak but significant interaction between Syntax and Semantics. Given the large P600 observed in response to ungrammatical sentences, we concluded that the cognitive resources reflected by this component were already used, so that no additional increase in amplitude was observed in the combined condition. In the present study we observed additive effects, which apparently contradict this interpretation. As it will become clear in the analyses of individual responses, the weak effect observed in L2 speakers can be explained by inter-individual variability in their ERP responses. Therefore, exploratory analyses of individual differences that we will discuss next turned out as a very helpful way of elucidating the real underlying patterns that are at play.

4.2. Individual differences

Inspection of individual data revealed some biases toward N400 or P600 profiles (Kim, Oines, & Miyake, 2018; Tanner et al., 2014; Tanner, McLaughlin, Herschensohn, & Osterhout, 2013), however in our data only lexical-semantic anomalies showed a significant correlation between the N400 and the P600 amplitudes. We propose two explanations for the discrepancy between our results and Tanner and colleagues' (see also van Hell & Abdollahi, 2017). First, they selected the *same* electrodes (at central and parietal sites) for both components and correlation measures. While central and parietal electrodes are representative for both the N400 and the P600 effects, it is expected that the two components highly impact each other, especially when considering their amplitude at the same site and adjacent time windows. Brouwer and Crocker (2017) point out that peaks observed in waveform-based component structures are only epiphenomena of latent components. That N400 and P600 effects correlate when measured at the same sites does not mean that the latent N400 and P600 truly correlate among participants. Our decision to quantify N400 and P600 effects at distinct electrodes and in non-adjacent time intervals may not have completely circumvented these issues, but should certainly have reduced the risk of finding a correlation that was simply due to the overlap of the two components. Second, Tanner et al (2013) focused on one dimension of language processing (morpho-syntactic agreement). We show that there is indeed a correlation between the N400 and P600 when investigating lexical-semantic anomalies alone, but not in conditions involving syntactic category violations.

In this study, we considered aspects that have been argued to be either sensitive to AoA or not (i.e. syntax vs semantics), and investigated the effects of AoA, proficiency, and exposure on these. Regression-based approaches have been argued to be appropriate to model continuous variables (van Hell & Tanner, 2012; Meulman et al., 2015), but we have seen that they are difficult to implement with a relatively large number of correlated variables. Using Random Forest allowed us to rank the most relevant predictors in order of importance, before using decision trees to establish their effects on ERPs.

Despite using an analysis package (i.e., *ranger*) that did not favour continuous over categorical predictor variables, factor ‘Group’ (L1 vs L2) always failed to emerge as a relevant factor in all conditions. Further, AoA and AoE were usually irrelevant with one exception: the P600 effect in the Combined condition. Even when considering the L2 group separately, AoA is systematically the least important variable, and AoE is always less important than daily usage and proficiency variables. In contrast, daily usage, immersion, and online performance were generally the most reliable factors explaining our data. Note that this finding does not deny AoA’s role in language learning, which has been well documented (e.g. Hernandez & Li, 2007; Steinhauer, 2014): there are more L2-learners in the lower clusters and more L1-speakers in the higher ones. Rather, it suggests that exposure and proficiency may be more appropriate measures accounting for variability in the data, and that splitting between groups may not be the most suitable way to analyze data. Thus, our results are inconsistent with the critical period hypothesis.

Percentage of daily usage was a determining factor accounting for the P600 response in the Syntax and Combined conditions: the more speakers were exposed to French, the larger their P600 response to these errors. Daily usage has not been the focus of L2 studies of sentence processing in adults; research on bilingual language acquisition and developmental language disorder, however, has shown that regular exposure (between 20-30 % daily) is essential to L2 command (Ojima, Nakata, & Kakigi, 2005) including morphosyntax (Elin Thordardottir, 2015; Marquis & Royle, 2019). Interestingly, we observe that participants with this amount of exposure or more elicit a P600 effect. It seems that for L1 and L2 speakers alike, regular exposure over a given threshold enhances the ability to recruit mechanisms that relate to rule-based, procedural processing (Ullman, 2004) and conflict monitoring (Vissers et al., 2008). This effect is not categorical: participants who receive the highest exposure (over 90%) show the largest P600 effect in response to 'pure' syntactic category violations. Daily usage was also an important variable accounting for the P600 effect in the Semantics condition, but inspection of a sample tree revealed no significant split.

As predicted by the convergence hypothesis, another central variable to account for both the N400 and the P600 effects was proficiency. Information provided by the cut-off values suggests different interpretations. The link between P600 effects and higher online performance in the Syntax condition suggests that this component is associated with more stringent categorization of inacceptable sentences. On the other hand, the N400 seems to be predicated on above chance levels in the online task: it could tentatively be interpreted either as a byproduct of participants paying attention to the task or as a lexical-semantic bottleneck (Universit, 2015). However, a bottleneck interpretation would be strongly supported by LexTALE effects, which we do not observe. In the Combined condition, LexTALE split the high-exposure cluster: participants with a better LexTALE score showed a smaller P600 effect. We are presently unsure what explanation could account for this surprising effect. Online performance effects could reflect attention effects in two (not mutually exclusive) ways: higher performance at the task could be associated with less noisy data in participants who are paying attention, or participants who score better on average display larger ERP effects in general.

Duration of French immersion accounts for variability in the Semantics condition. It may reflect how the size of the lexicon increases with experience as seen in ERP studies with children (Ojima et al., 2005) and behavioral studies with aging adults (Cohen-Shikora & Balota, 2016; Robert & Rico Duarte, 2016; Royle, Steinhauer, Dessureault, Herbay, & Brambati, 2019). Digit span score was the most important predictor in that condition, and tree inspection showed that participants with a lower score on the working memory digit span task (below 9) did not elicit any N400 effect in response to lexical-semantic anomalies, suggesting that participants may need to recall context to detect anomalies in the experimental sentence. More surprisingly, participants with a high digit-span score (above 12) elicited a negativity instead of a P600 in response to the Combined condition. The relationship between working memory load and sustained negativities using similar experimental materials in the auditory modality is currently under evaluation (Fromont, Royle & Steinhauer, in preparation) and may shed light on this effect.

Inspection if individual differences using Random Forests and trees helps us understand the seemingly small P600 (or absent in the case of syntactic category violations) effects in the L2 group. Data partitioning shows that participants with less exposure who are L2ers do not display any P600 effect, while L2ers who are more exposed do. Further, some participants even display *opposite* effects in the late time-window: subgroups show a relative parietal negativity (at sites the P600 is measured) in the syntactic category violation condition and a large negativity in the Combined condition. These negativities likely attenuated or canceled the group effects, suggesting that Random forests are an appropriate method to identify the variables that truly contribute to the variability in our data (i.e. daily usage, performance, and digit span score, rather than group).

5. Conclusion

Our results demonstrate that both native speakers and proficient late L2-learners process syntactic information in parallel with lexical-semantic information. At a group level, native speakers elicited a biphasic N400-P600 response to syntactic category violations – and no ELAN or LAN – while L2-learners only displayed a significant N400 effect. However, when investigating interindividual variability and the relative importance of predictor variables related to AoA, L2 exposure, and proficiency, we found that daily usage, immersion, and proficiency – not AoA – were the most important predictors for the observed ERP components. Both L1- and L2-speakers display larger N400 and P600 responses to syntactic errors as their daily usage, immersion, and proficiency increase. This evidence lends further support to the convergence hypothesis, and suggests that higher language exposure and proficiency are associated with both memory-based heuristics and rule-based processes.

General Discussion

1. A critical evaluation of existing models in light of the present results

1.1. Friederici's "syntax-first" model

The "syntax-first" model (Friederici, 2002; 2011; 2012) predicts an ELAN followed by a P600 in response to syntactic category violations, and no N400, even for combined anomalies. The ERP data presented here clearly demonstrate that ERPs elicited by syntactic category violations and grammatical targets are perfectly aligned until 300 ms. The absence of any early ERP effect provides strong evidence that there is no primacy of syntactic category identification in sentence processing. Further, an N400 effect was observed in response to syntactic category violations, and an even larger N400 was elicited by the combined SC and lexical-semantic anomalies. This evidence shows that there is no blocking of lexical-semantic processing in the presence of SC errors. In sum, the postulate that the parser is, at first, blind to any information except for SCs, and that lexical-semantic processing is contingent upon SC identification does not hold. We also provided evidence that lexical-semantic anomalies elicited a small P600: the observed interactive effects of our SC and lexical-semantic manipulations on this component suggest that the same resources are used up to process both types of anomalies. This is inconsistent with a view of the P600 as an index of structural reanalysis, as postulated by the "syntax-first" model. As most of the model's predictions were not supported by our data, and considering the methodological concerns outlined by S&D2012, as well as evidence from studies that did not observe an ELAN (Bowden et al., 2013; Nickels, 2016; White et al., 2006; Yang et al., 2015; Zhang et al., 2013), the "syntax-first" model in its current form– strictly serial and modular – should be abandoned.

However, our data are compatible with certain aspects of serial models. One of them is that a top-down parser commits to a specific syntactic structure in advance – before getting to the lexical item that fulfills it (e.g., Staub & Clifton, 2006). A related idea is that local relationships have a privileged status over higher-order ones, due to hard-wired preferred syntactic structures. (e.g., Frazier & Fodor, 1978). Alternatively, the primacy of the syntactic analysis can be motivated by computational limitations (e.g., memory: Gibson, 1998), and by the fact that some categories are simply easier to predict: for example, in head-initial languages, heads highly constrain the complement that will eventually follow, and have lower constraints on optional adjuncts[15]. We will further discuss this interpretation in the assessment of predictive approaches and in our own proposal; but first, we will evaluate the alternative possibility of parallel, interactive syntactic and semantic streams.

1.2. Parallel approaches

The MUC model (Hagoort, 2005, 2016) hypothesizes that sentence processing relies on separate semantic and syntactic streams that can interact at any point. In this framework, the N400 reflects semantic unification (*not* retrieval nor even integration: Hagoort et al., 2009) and the (L)AN-P600 reflect syntactic unification. The MUC model explains our data better than Friederici's account, since it postulates that syntactic and semantic information are processed in parallel. However, this model needs to account for cases where a "semantic N400" is observed in response to (at least seemingly) syntactic errors (Nickels, 2016), and where a "syntactic P600" is found in response to semantic anomalies.

[15] Investigating cross-linguistic differences in predictive processes, for example by comparing head-initial and head-final languages, may help clarify this issue.

In previous studies supporting the MUC model, Hagoort (2003b), and Van den Brink and Hagoort (2004) observed a (L)AN in response to syntactic category violations (*Het vrouwtje veegde de vloer met een oude *kliederde gemaakt van twijgen*[16]): this effect was interpreted as a failure to bind two constituents (e.g., the adjective 'old' with the verb 'messed'). However, the syntactic category violations in the present studies elicited N400 effects, presumably because the violations were created by manipulating the word-stem rather than affixes. Under the hypothesis that the N400 reflects semantic unification, a "loser takes all" explanation would not hold, since syntactic and semantic cues do not compete (contrary to the case presented above for argument role reversals). A plausible interpretation of our data would be that the semantic stream took over and that participants were having difficulty achieving a coherent semantic representation for the presented sentences: although it was not explicitly predicted by Hagoort, this interpretation seems generally consistent with the MUC. The subsequent P600 effect in response to syntactic category violations would then index a failure to achieve any hierarchical representation of the sentence structure. This model, however, cannot explain the presence of a (small but significant) P600 effect in response to "pure" lexical-semantic anomalies.

Turning to the very nature of these unification semantic processes, it is unclear how the MUC model would interpret the additive effects of lexical-semantic anomalies and syntactic category violations on the N400 component. The definition proposed by Hagoort and colleagues for semantic unification suggests that every word is integrated incrementally within one single process:

---

[16] 'The woman swept the floor with an old *messed made of twigs'.

> "Semantic unification refers to the integration of word meaning into an unfolding representation of the preceding context. This is more than the concatenation of individual word meanings […]. In the interaction with the preceding sentence or discourse context, the appropriate meaning is selected or constructed, so that a coherent interpretation results." (Hagoort et al., 2009, Chap. 56, p. 1)

In this context, priming should not really have an additive effect on the N400 effect when there is a syntactic category violation, since a coherent representation cannot be built on the basis of the syntactic category violation alone in any case. However, one could consider different layers of semantic unification: in the present experiment, one layer would be at the level of priming between the context and a target, and another at the level of local relationships between words of the experimental sentence. Adding multiple layers of unification would allow the MUC model to account for this data. Finally, as the P3a component normally reflects surprisal and reallocation of attention, it is normally expected after lexical retrieval or prediction-error mechanisms. Since the N400 that precedes it already reflects unification processes, the MUC model would probably interpret the frontal positivity in the "pure" syntactic category violation condition as syntax-specific rather than a domain-general P3a – unlike what has been argued in Chapter 2.

1.3. Predictive approaches

Lau and collaborators (2006) argue that strong structural predictability speeds up the retrieval of SCs and promote ELANs. Our present experiment uses contexts that (i) are syntactically unambiguous: transitive and control verbs unambiguously select noun and verb complements, respectively, and (ii) which promote a rapid diagnosis on the target word (the presence of clitic / determiner *le-la-les* introduces a delay between the main verb and the target allowing for clear predictions to emerge). The degree of high constraint for specific syntactic categories is further evidence by our cloze test results, which show that in over 40 participants not one inserted the 'wrong' syntactic category after the main verb. Yet, there was no early activation in the form of an ELAN, as predicted by Lau et al. Pre-activation of target words is a very important argument for predictive approaches, but there is very little empirical evidence that the early effects genuinely reflect these predictive processes (DeLong et al.'s famous study on lexical-semantic pre-activation has also not been replicated, cf. Nieuwland et al., 2018).

Nevertheless, predictive approaches can survive without the ELAN component. In fact, most predictive views focus on the N400 component as an indirect index of predictive processes. Under this view, the N400 effect mostly reflects retrieval (Lau et al., 2008): if a word has been successfully predicted, then it is easier to retrieve and its N400 amplitude is lower. Our finding that cloze probability modulates the N400 effect in the primed condition regardless of syntax is completely in line with this view. While some accounts distinguish between prediction and priming (e.g., Otten, Nieuwland, & Van Berkum, 2007), others do not exclude the possibility that priming mechanisms are part of predictive processes if they result in pre-activation of a set of lexical representations (Bornkessel-Schlesewsky & Schlesewsky, 2019; Chow, Momma, et al., 2016; but see Lau, Holcomb, & Kuperberg, 2013). ERP studies that explicitly evaluate prediction have not really focused on syntactic categories *per se* (excepting Lau and colleagues' 2006 study discussed above). One possible interpretation stemming from our data is that in a highly constraining syntactic context where a complement noun or verb is expected, encounters with a lexical item of the wrong category elicit an N400 effect, indexing lexical retrieval difficulties. Several possibilities arise. First, we can hypothesize that local grammatical constraints facilitate retrieval by boosting the pre-activation of certain lexical candidates that belong to the right SC.

A behavioral multi-experiment study on the effects of syntactic constraints on lexical search (Wright & Garrett, 1984) may further inform the 'lexical boosting' hypothesis. The authors used reaction times measured during a lexical decision task on semantically unexpected target words. Similarly to our study, the authors manipulated the syntactic category of lexical items appearing after control or transitive verbs in English (*If your bicycle is stolen, you must formulate/*batteries./For now, the happy family lives with batteries/*formulate.)* and showed that syntactic category violations elicited longer reaction times than their syntactically correct (but still incongruous) equivalent. Similar results were obtained when exclusively manipulating the affix marker for SC (e.g., *translates vs. 'translation*), but not for non-words that bore the same affixes (e.g., *fablorates' vs. 'fabloration*). Further, these effects were only observed when the target words occupied an 'obligatory' phrasal position (i.e., as a head rather than an adjunct). The authors suggest that taken together, the observed effects may be due to either (i) a confirmation procedure (i.e. checking whether the structural constraints are satisfied by a given lexical item) or (ii) lexical search facilitation (i.e. the required syntactic category is a point of entry to the lexicon). If (ii) were true, we would find syntactic category violation effects prior to lexical-semantic effects, but the two appear in parallel. Hypothesis (i) is possible, but seems to suggest that the confirmation can only happen after the lexical item has already been retrieved, before it is transferred to working memory (Steinhauer et al., 2017); how it would be compatible with a retrieval view of the N400 that is advocated by predictive approaches is unclear.

Under the assumption that retrieval mechanisms influenced by prediction are fast and incremental, it is not very surprising that they would work in parallel. On the one hand, participants are keeping lexical items from the context sentence (e.g. *hockey*) activated using their working memory until the end of the second sentence and, by doing so, allow for priming effects to occur (explaining the N400 elicited in the lexical-semantic condition). On the other, they make predictions about each incoming word using local information, which has a privileged status because it is a constraining context and they do not need to engage as much memory resources. These predictions pre-activate another set of words that are plausible complements for the verb (i.e. the right syntactic category and they make sense).

Integration of this information at the interface between syntax and semantics, however, may take more time and be costly in terms of cognitive resources (which is also what was observed for the "semantic P600", see Chow, Smith, Lau, & Phillips, 2016). Note also that predictive approaches do not provide an extensive functional interpretation of P600 effects.

The data presented in this thesis are coherent with a sentence-processing proposal where, first, retrieval is facilitated by prediction mechanisms, which may still be distinct depending on the linguistic information that is provided (structural cues, priming), and, later, qualitatively distinct processes encompassing a range of mechanisms (including structural analysis, task effects and conflict monitoring) are reflected by the P600 (Chow, Smith, et al., 2016). However, it is unclear how this interpretation is tenable with data showing that some individuals tend to display either an N400 or a P600 dominance to grammatical errors (Tanner, 2019).

2. Inter-individual variability and underlying cognitive processes

Studies focusing on the individual ERP responses elicited by agreement errors (Tanner & Van Hell, 2014), argument-structure errors (Kim et al., 2018), syntactic category violations and lexical-semantic errors (Nickels, 2016) have demonstrated that there exists a negative correlation between the N400 and the P600 effect. This tradeoff has been described as "replicable and robust" (Tanner, 2019), and it has been further demonstrated that individuals hold a similar "ERP profile" across distinct experimental manipulations (Tanner & Van Hell, 2014). One interpretation of these results is that the N400 and the P600 indeed subserve cognitively distinct processes, but that individuals tend to rely more strongly on either one of these processes during sentence comprehension. This pattern is not observed in this thesis[17]: most participants elicited an N400 for both syntactic category violations and lexico-semantic anomalies, and the magnitude of this effect did not correlate across conditions. The magnitude of the P600 effect, however, did correlate across conditions, suggesting that the same underlying cognitive resources were used by given participants regardless of the anomaly.

The convergence theory (Green, 2003; Steinhauer et al., 2009; Steinhauer, 2014) predicts that ERP profiles of L2 participants converge toward L1 patterns with increasing L2 proficiency. While it does *not* predict that the P600 effect is a marker of nativelikeness (as suggested by Tanner et al), it has not provided a characterization of L1 speakers variability either (which may have triggered Tanner's criticism). The study presented in Chapter 3 explores inter-individual variability from data collected with 81 participants, without *a priori* distinctions between L1 and L2 status. Random forests revealed that daily usage and

[17] In §4 of the present discussion, a methodological explanation for this discrepancy is presented.

performance were very important predictor for both N400 and P600 components in L1 and L2 speakers. Specifically, it suggests that participants who use French above 20-30% of the time in their daily life elicit native-like P600 effects in response to ungrammatical sentences. Participants who are less exposed still show small effects with sufficient proficiency. These results are in line with the convergence theory, and – importantly – they show that similar characterizations are true of L1 and L2 processing.

We are not aware of any ERP study that directly investigates the effect of the amount of daily usage on sentence processing in bilingual adults. However, our results are in line with behavioral studies on bilingual children (Barriere, 2010; Elin Thordardottir, 2015; Marquis & Royle, 2019; Ojima et al., 2005). Even though the amount of daily usage did not overlap between groups, which was a potential bias, the conditional inference trees split the data in a way that L2 speakers were present in higher clusters where a P600 was observed. However, future research could investigate this effect further by matching L1 and L2 speakers with respect to the amount of daily linguistic exposure they receive.

In the present studies, online performance at the task is used as a proxy for proficiency (Nickels & Steinhauer, 2018; Steinhauer et al., 2009). However, this variable may not quite reflect the same characteristics in L1 and L2 speakers: while some L2 learners may not notice a sentence is ungrammatical because they are not proficient enough, L1 speakers should *know* all syntactic category violations are unacceptable in French. At this point, one can only speculate about what being a "high performer" means (especially for native speakers): this notion is most probably conflated with the allocation of attentional resources, working memory abilities (e.g., Mecklinger, Schriefers, Steinhauer, & Friederici, 1995) and general motivation (Nickels, 2016). Interestingly, there seems to be an apparent contradiction between the association, in the literature, of the P600 with slow and costly processes triggered by sentence reanalysis, and the observation that the P600 effect is larger for individuals who perform better in our data. Why would proficient individuals rely on these costly processes in addition to (or instead of) predictive mechanisms which appear to be faster and more efficient? One possibility is that "P600ers" can "afford to wait", because they have plenty of cognitive resources (i.e., attentional or working memory resources, Mecklinger et al., 1995) in conjunction with, in the case of L1 speakers, a complete command of grammatical rules of the target language, to adjudicate between different representations, and are thus able to work at the syntax-semantics interface during sentence comprehension.

3. A preliminary account of sentence processing in L1 and L2

As far as sentence processing accounts are concerned, the two best candidates to explain the present data make drastically different predictions. The MUC model relies heavily on the distinction between syntax and semantics, while considering that both the N400 and the P600 reflect unification processes ("the P600 is the syntactic equivalent of the N400", Hagoort et al., 2009, Chap 56, p. 6). On the other hand, recent "predictive" approaches, as well as work on inter-individual variability, tend to distinguish between fast retrieval / predictive processes and later "combinatorial" processes (Osterhout et al., 2006; Tanner et al., 2013), but blur the distinction between semantics and syntax. In fact, some accounts such as the retrieval–integration model (Brouwer, Fitz, & Hoeks, 2012; Brouwer & Hoeks, 2013) make the syntax-semantics distinction simply irrelevant. The data presented in this thesis does *not* suggest that the distinction between semantic and syntactic processes as reflected by ERP components should be abandoned: syntactic category violations *still* modulate the P600 effect over and above lexical-semantic anomalies, and priming modulates the N400 in addition to SC manipulations.

A preliminary reevaluation of the time-course of sentence processing is outlined in Figure 15. This model takes into account the evidence available in: (i) SC studies that use a sound methodological design (including the present one); (ii) agreement studies with a particular focus on the debate around the LAN-N400 effects, and, to a much lesser extent; (iii) argument-structure studies that are not the focus of the present thesis, but have been greatly debated in recent years, and have been the source of many proposals on predictive approaches (Chow, Smith, et al., 2016; Kim et al., 2016; Kuperberg et al., 2006).

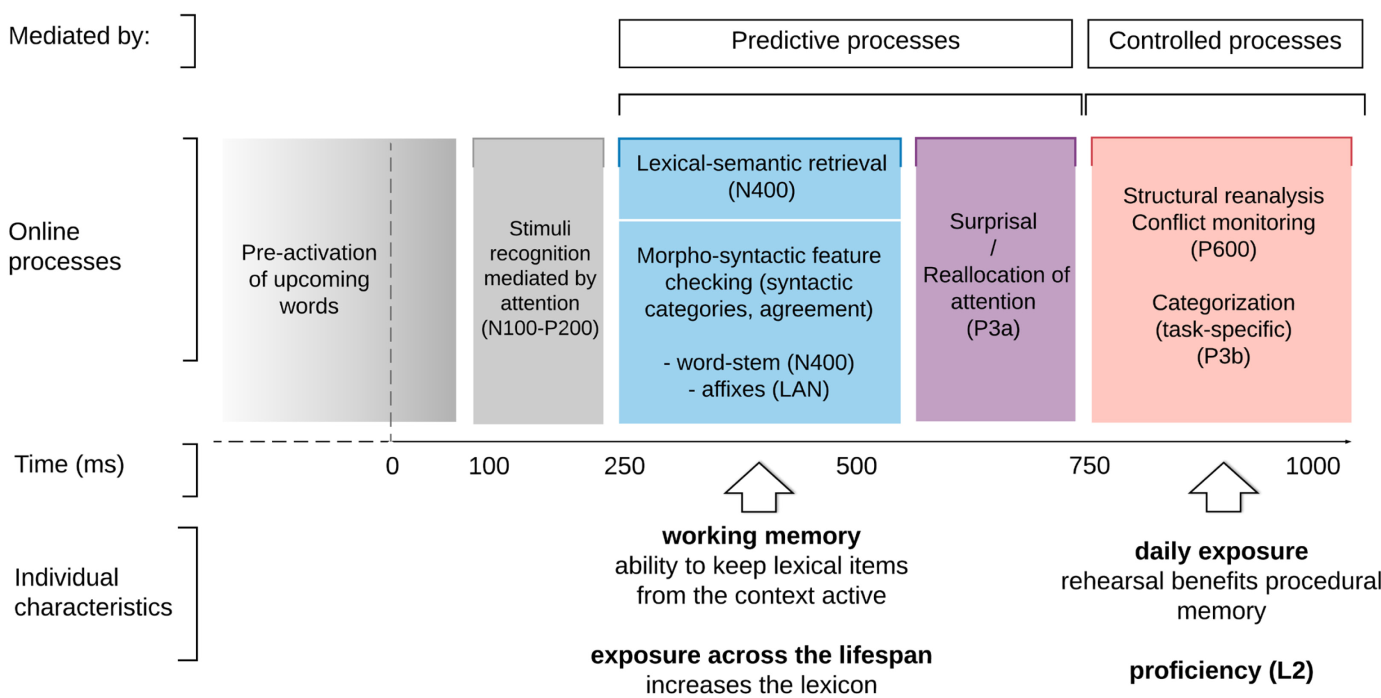


Figure 15. A reevaluation of the time-course of sentence processing.

As hypothesized by predictive approaches, this model assumes that specific lexical sets can be pre-activated before the target word appears. Early stages of word recognition are not the focus of this account. In the visual–reading modality, we consider that early processes can be described along the lines of Sereno and Rayner's (2003) time course, where visual information takes about 60 ms to reach the visual cortex, leading to a shift in attention that is also sensitive to lexical properties (such as frequency). After word recognition, there is a distinction between mechanisms that are mediated by predictive processes (i.e. via pre-activation of lexical sets), and controlled processes. In the earlier stages of retrieval (starting ~250 ms) the activation of a set of lexical-semantic entities occurs (Lau et al., 2008). At the same time, and potentially in parallel, cohorts with the morpho-syntactic features of the expected word (syntactic categories, agreement) are activated. The parser uses the available information, which may come in the form of affixes[18] (typically eliciting LAN effects) or word-stem information (eliciting N400 effects; Molinaro et al., 2015) to evaluate candidates. To our knowledge, although the "syntax-first" model would have suggested that SC information also constrains agreement processing (Rossi et al., 2005), no study has addressed the interplay between agreement and SC identification using a sound experimental design. For now, we speculate that these processes can interact. However, under the hypothesis that retrieval mechanisms are fast and are mediated by predictive processes, which themselves depend on the amount and type of information provided by the context, they may not need to interact at first. Also note that our data does not allow us to determine to what extent lexical-semantic relationship between adjacent constituent may affect the N400 in response to syntactic category violations. After 550 ms, depending on the experimental context, surprising stimuli may trigger an additional allocation of attentional resources (P3a, Kasparian et al.,

[18] Although at least some violations of derivational affixes have elicited LAN effects – see Leminen, Smolka, Duñabeitia, and Pliatsikas (2018), for a review – this proposal focuses on inflectional morphology, not derivational

2017). Processes that are controlled include conflict monitoring (Vissers et al., 2008), structural reanalysis (e.g., Osterhout & Holcomb, 1992), and categorization (Sassenhagen et al., 2014), are all reflected by the P600 / P3b, and are presumably highly interactive (in fact, it is difficult to tease these effects apart). Individual factors such as high working memory span promote priming effects by keeping relevant lexical items active in memory, reflected by the N400, and exposure and robust grammar command promote at least certain processes that are reflected by the P600.

4. Implications for ERP research

The present study provides empirical support for the methodological arguments made by S&D2012, who demonstrate that at least some context manipulation designs lead to artefacts which, as in the case of Friederici's paradigm, may be misinterpreted as genuine ERP effects. Despite this concern, some paradigms continue to fail at keeping contexts constant (e.g. Nichols & Joanisse, 2019): in those cases, researchers should at least show a longer baseline interval to support the absence of artefact-inducing upstream effects, and explicitly consider the possibility of using a distant baseline (e.g., Friederici et al., 1999; Nickels 2016). Target manipulation designs, in SC studies in particular, may also create biases: for example, participants could associate a certain SC with a specific condition (e.g., nouns with syntactic category violations and verbs with correct controls in Yang et al., 2015). The solution to these issues is to create a balanced design. Note that the advantages of using such designs are not limited to SC studies: they have been demonstrated in agreement studies as well (e.g., to investigate the effects of omission vs. commission of the third person singular verb marker *-s* in English: Dube, Kung, Peter, Brock, & Demuth, 2016, or verb marking in French: Courteau et al., 2019). Using a balanced design does not prevent one from looking at specific sub-conditions under certain circumstances – after validating that different contexts do not lead to any baseline differences. In the present studies, the ERPs elicited by the correct sub-conditions were different, which unfortunately prevented us from interpreting specific effects related to either noun or verb targets. However, had we not used a balanced design, our interpretation would have been blind to this possible bias.

Exploring inter-individual variability was not a primary objective of this study. However, the large number of items per condition ($n = 80$) and of participants in both L1 and L2 groups ($n = 81$) as well as our carefully controlled design contributed to reducing noise and made exploration of individual responses possible. Tanner's argument about the LAN (2015) being a by-product of component overlap between N400 and P600 effects on the grand average was extremely relevant to this study, since one could have predicted a LAN effect in response to syntactic category violations in French. However, it is rather surprising that, despite their concerns about the effects of component overlap on the "LAN effect", Tanner and collaborators ignore the possibility of component overlap when examining the correlation between the N400 and the P600 components for each participant. They typically use adjacent time-windows and similar ROIs even though the N400 and the P600 spatiotemporally overlap. It is therefore extremely difficult to examine their relation independently (Brouwer & Crocker, 2017); non-adjacent time-windows and different ROIs may be more informative of the real relationship between the N400 and the P600. In this study, we "replicated" the "tradeoff effect" between N400 and P600 (Kim et al., 2018) when using adjacent time-windows and a ROI that included both C and P electrodes, but *did not* when using different ones to minimize artefacts due to component overlap. This suggests that the tradeoff hypothesis should be taken with caution, and that analyses that consider the whole biphasic waveform without taking topographical differences into account (Meulman et al., 2015) should be avoided.

Nevertheless, inter-individual variability is real, and so may be at least *some* response dominance profiles. It is to be expected that research will increasingly focus on this variability, and one long-term objective is of course to be able to group individuals together based on specific predictors (Tanner, 2019). This objective is important in the context of L2 processing, since we seek to understand what circumstances favor successful language learning, but it is also instrumental for a better understanding of the time-course of sentence processing in general. In the present study we explore random forests as a complementary tool to linear regressions to investigate the importance of individual predictors. Although there certainly is room for improvement, this analysis has led to encouraging results that will hopefully motivate other ERP researchers to harness machine learning tools as new ways to analyze their data.

5. Limitations and future directions

Considering the limitations of this study should also provide some guidelines for future research on the time-course of phrase-structure building. First, as mentioned in the previous section, category-specific effects are absent from the analyses presented in chapters 2 and 3 because it was not possible to split between sub-conditions. Second, as the word-stem was not kept constant across conditions, it is difficult to tease apart "pure" SC identification processes from lexical-semantic ones.

The main verbs were matched and selected their ability to take either noun or verb complements. We saw that some of the control verbs could be coerced into taking noun complements (either in complement or adjunct position, $n = 7$). In addition, in principle, pre-nominal adjectives could be placed after the transitive verbs. We did not expect our participants to predict this specific outcome because (i) there are relatively few adjectives that are pre-nominal in French, (ii) our experimental design was repetitive (which, however, means it does not quite replicate ecological conditions), and (iii) adjectives are generally superfluous unless they are used to distinguish between two possible referents (Valois & Royle, 2009). On the other hand, note that in the cloze test, participants were asked to complete the full sentences and all of them placed nouns after transitive verbs and verbs after control verbs. In the end, we conclude that the contexts in our experiment were indeed quite constraining for word category. Despite this, we still did not observe an ELAN.

On the topic of inter-individual variability, some additional information about participants would have been useful in analyzing the data. First, partly because of the socio-linguistic situation in Montreal, exposure was confounded with AoA, and this makes it difficult to tease apart their relative effects on L2 processing. Second, after considering the effects of working memory on individual ERP responses, it appears in hindsight that it would have been informative to collect additional measures of working memory such as the reading-span task (Just & Carpenter, 1992). In an ERP auditory study using the same stimuli (Fromont, Royle, & Steinhauer, in preparation), this measure was included.

As “reliable” ERP effects in language studies increasingly seem to be restrained to the N400 and the P600, there are questions about the underlying mechanisms they are supposed to reflect: are they multiple, and if so, do they interact? Exploring the additivity of ERP effects may help answer those questions. To do this, stronger anomalies should be created to promote larger effects and to ensure that if indeed similar cognitive resources underlie two different processes (say, SC identification and lexical retrieval), they would have to compete.

Finally, in order to close the gap between the reading and the auditory modalities, and because the ELAN itself has sometimes been argued to be specific to the auditory modality, the same materials will have to be evaluated in that modality as well. This has been done, and the design used includes an acoustic manipulation on the determiner / clitic which may help listeners predict the upcoming SVC, and therefore inform the predictive approach (Fromont, Royle, & Steinhauer in preparation).

Conclusion

The studies presented in this dissertation are among the few existing ones to empirically address the shortcomings of previous ELAN studies that supported the "syntax-first" model of sentence processing. Specifcially, the experiments are the very first to systematically manipulate SCs and lexical-semantic priming while keeping the design completely balanced. The careful elaboration of the experimental stimuli provided some of the cleanest ERP data obtained in electrophysiological studies of morpho-syntactic processes to date.

The results obtained in L1 and L2 speakers of French, showing N400 effects in response to SC violations, consitute strong evidence against serial approaches of sentence processing. This work also clarified neurocognitive processes underlying phrase-structure building where existing studies (even those published after S&D2012) have conflicting results.

In addition, the relatively large sample of L1 and L2 participants allows us to address some issues related to inter-individual variability that are currently under debate. Specifically, our data suggest that the underlying mechanisms reflected by the N400 and P600 effects *can* both be at play during sentence processing, contrary to what has been suggested (Kim et al., 2018; Tanner et al., 2014, 2013; Tanner & Van Hell, 2014).

These studies, along with consideration of existing literature on online SC identification, lexical-semantic processing and prediction, and agreement processing, led to a proposal characterizing the dynamics of sentence comprehensiom that will hopefully contribute to inspire future research in native and second language processing.

References

Babin, J.-P. (1998). Caractéristiques des items lexicaux affectant leur reconnaissance. In *Lexique mental et morphologie lexicale* (pp. 19–56). Berne: Peter Lang.

Baggio, G., & Hagoort, P. (2011). The balance between memory and unification in semantics: A dynamic account of the N400. *Language and Cognitive Processes*, *26*(9), 1338–1367. DOI:10.1080/01690965.2010.542671

Barber, H., & Carreiras, M. (2005). Grammatical gender and number agreement in Spanish: An ERP comparison. *Journal of Cognitive Neuroscience*, *17*(1), 137–153. DOI:10.1162/0898929052880101

Barriere, I. (2010). The vitality of Yiddish among Hasidic infants and toddlers in a low SES preschool in Brooklyn. *Yiddish — A Jewish National Language at 100, Proceedings of Czernowitz Yiddish Language 2008 International Centenary Conference*, (January 2010), 170–196.

Basciano, B. (2017). Word classes, Modern. In R. Sybesma, W. Behr, Y. Gu, Z. Handel, C.-T. J. Huang, & J. Myers (Eds.), *Encyclopedia of Chinese language and linguistics* (pp. 554–566). Leiden: Brill.

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1). DOI:10.18637/jss.v067.i01

Bauer, D. J., & Sterba, S. K. (2011). Fitting multilevel models with ordinal outcomes: Performance of alternative specifications and methods of estimation. *Psychological Methods*, *16*(4), 373–390. DOI:10.1037/a0025813

Bernal, S., Dehaene-Lambertz, G., Millotte, S., & Christophe, A. (2010). Two-year-olds compute syntactic structure on-line. *Developmental Science*, *13*(1), 69–76. DOI:10.1111/j.1467-7687.2009.00865.x

Birdsong, D. (2006). Age and second language acquisition and processing: a selective overview. *Language Learning*, *56*(s1), 9–49. DOI:10.1111/j.1467-9922.2006.00353.x

Birdsong, D. (2018). Plasticity, variability and age in second language acquisition and bilingualism. *Frontiers in Psychology*, *9*, 81. DOI:10.3389/fpsyg.2018.00081

Bokhari, F. S. (2015). *Investigating the neurocognitive mechanisms underlying truth-conditional and logical semantic aspects of sentence processing: an ERP study*. McGill University.

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2008). An alternative perspective on "Semantic P600" effects in language comprehension. *Brain Research Reviews*, *59*(1), 55–73. DOI:10.1016/j.brainresrev.2008.05.003

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2019). Towards a neurobiologically plausible model of language-related, negative event-related potentials. *Frontiers in Psychology*, *10*, 298. DOI:10.3389/fpsyg.2019.00298

Bornkessel, I., Fiebach, C. J., & Friederici, A. D. (2004). On the cost of syntactic ambiguity in human language comprehension: An individual differences approach. *Cognitive Brain*

Research, *21*(1), 11–21. DOI:10.1016/j.cogbrainres.2004.05.007

Bornkessel, I., & Schlesewsky, M. (2006). The extended argument dependency model: A neurocognitive approach to sentence comprehension across languages. *Psychological Review*, *113*(4), 787–821. DOI:10.1037/0033-295X.113.4.787

Bowden, H. W., Steinhauer, K., Sanz, C., & Ullman, M. T. (2013). Native-like brain processing of syntax can be attained by university foreign language learners. *Neuropsychologia*, *51*, 2492–2511. DOI:10.1016/j.neuropsychologia.2013.09.004

Breiman, L. (2001). Random forests. *Machine Learning*, 5–32. DOI:10.1023/A:1010933404324

Brouwer, H., & Crocker, M. W. (2017). On the proper treatment of the N400 and P600 in language comprehension. *Frontiers in Psychology*, *8*, 1327. DOI:10.3389/fpsyg.2017.01327

Brouwer, H., Fitz, H., & Hoeks, J. (2012). Getting real about Semantic Illusions: rethinking the functional role of the P600 in language comprehension. *Brain Research*, *1446*, 127–143. DOI:10.1016/j.brainres.2012.01.055

Brouwer, H., & Hoeks, J. C. J. (2013). A time and place for language comprehension: mapping the N400 and the P600 to a minimal cortical network. *Frontiers in Human Neuroscience*, *7*(November), 1–12. DOI:10.3389/fnhum.2013.00758

Brusini, P., Dehaene-Lambertz, G., Dutat, M., Goffinet, F., & Christophe, A. (2016). ERP evidence for on-line syntactic computations in 2-year-olds. *Developmental Cognitive Neuroscience*, *19*, 164–173. DOI:10.1016/j.dcn.2016.02.009

Brusini, P., Dehaene-Lambertz, G., van Heugten, M., de Carvalho, A., Goffinet, F., Fiévet, A. C., & Christophe, A. (2017). Ambiguous function words do not prevent 18-month-olds from building accurate syntactic category expectations: An ERP study. *Neuropsychologia*, *98*, 4–12. DOI:10.1016/j.neuropsychologia.2016.08.015

Brysbaert, M. (2013). Lextale_FR: a fast, free, and efficient test to measure language proficiency in French. *Psychologica Belgica*, *531*, 23–37. DOI:10.5334/pb-53-1-23

Caffarra, S., Mendoza, M., & Davidson, D. (2019). Is the LAN effect in morphosyntactic processing an ERP artifact? *Brain and Language*, *191*, 9–16. DOI:10.1016/j.bandl.2019.01.003

Caffarra, S., Molinaro, N., Davidson, D., & Carreiras, M. (2015). Second language syntactic processing revealed through event-related potentials: an empirical review. *Neuroscience and Biobehavioral Reviews*, *51*, 31–47. DOI:10.1016/j.neubiorev.2015.01.010

Carreiras, M., Perea, M., Vergara, M., & Pollatsek, A. (2009). The time course of orthography and phonology: ERP correlates of masked priming effects in Spanish. *Psychophysiology*, *46*(5), 1113–1122. DOI:10.1111/j.1469-8986.2009.00844.x

Carreiras, M., Salillas, E., & Barber, H. (2004). Event-related potentials elicited during parsing of ambiguous relative clauses in Spanish. *Cognitive Brain Research*, *20*, 98–105. DOI:10.1016/j.cogbrainres.2004.01.009

Chow, W. Y., Lago, S., Barrios, S., Parker, D., Morini, G., & Lau, E. (2014). Additive effects of repetition and predictability during comprehension: Evidence from event-related potentials. *PLoS ONE*, *9*(6), e99199. DOI:10.1371/journal.pone.0099199

Chow, W. Y., Lau, E., Wang, S., & Phillips, C. (2018). Wait a second! delayed impact of argument roles on on-line verb prediction. *Language, Cognition and Neuroscience*, *0*(0), 1–26. DOI:10.1080/23273798.2018.1427878

Chow, W. Y., Momma, S., Smith, C., Lau, E., & Phillips, C. (2016). Prediction as memory retrieval: timing and mechanisms. *Language, Cognition and Neuroscience*, *31*(5), 617–627. DOI:10.1080/23273798.2016.1160135

Chow, W. Y., Smith, C., Lau, E., & Phillips, C. (2016). A "bag-of-arguments" mechanism for initial verb predictions. *Language, Cognition and Neuroscience*, *31*(5), 577–596. DOI:10.1080/23273798.2015.1066832

Christensen, R. H. B. (2015). A tutorial on fitting cumulative link mixed models with clmm2 from the ordinal package. Retrieved from https://cran.r-project.org/web/packages/ordinal/vignettes/clmm2_tutorial.pdf

Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistcs*, *27*, 3–42. DOI:10.1017.S0142716406060024

Cohen-Shikora, E. R., & Balota, D. A. (2016). Visual word recognition across the adult lifespan. *Psychology of Aging*, *31*(5), 488–502. DOI:doi:10.1037/pag0000100

Coulson, S., King, J. W., & Kutas, M. (1998). Expect the unexpected: Event-related brain potentials to morphosyntactic violations. *Language and Cognitive Processes*, 13(1), 21–58.

Courteau, É., Martignetti, L., Steinhauer, K., & Royle, P. (2019). Eliciting ERP components of morphosyntactic agreement mismatches in grammatical sentences. *Frontiers in Psychology*, *10*, 1152. DOI:10.3389/fpsyg.2019.01152

de Aguiar, V., Bastiaanse, R., & Miceli, G. (2016). Improving production of treated and untreated verbs in aphasia : a meta-analysis, *10*, 468. DOI:10.3389/fnhum.2016.00468

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience*, *8*(8), 1117–1121. DOI:10.1038/nn1504

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics. *Journal of Neuroscience Methods*, *134*, 9–21.

Diaz, M. T., & Swaab, T. Y. (2007). Electrophysiological differentiation of phonological and semantic integration in word and sentence contexts. *Brain Research*, *1146*(1), 85–100. DOI:10.1016/j.brainres.2006.07.034

Díaz-Uriarte, R., & Alvarez de Andrés, S. (2006). Gene selection and classification of microarray data using random forest. *BMC Bioinformatics*, 7, 1–13. DOI:10.1186/1471-2105-7-3

Diedenhofen, B., & Musch, J. (2015). Cocor: A comprehensive solution for the statistical comparison of correlations. *PLoS ONE*, *10*(4), e0121945. DOI:10.1371/journal.pone.0121945

Dikker, S., Rabagliati, H., & Pylkkänen, L. (2009). Sensitivity to syntax in visual cortex. *Cognition*, *110*(3), 293–321. DOI:10.1016/j.cognition.2008.09.008

Dube, S., Kung, C., Peter, V., Brock, J., & Demuth, K. (2016). Effects of type of agreement violation and utterance position on the auditory processing of subject-verb agreement: An ERP study. *Frontiers in Psychology*, *7*(AUG), 1–18. DOI:10.3389/fpsyg.2016.01276

Elin Thordardottir. (2015). The relationship between bilingual exposure and morphosyntactic development. *International Journal of Speech-Language Pathology*, *17*(2), 97–114. DOI:10.3109/17549507.2014.923509

Featherston, S., Gross, S., Münte, T., & Clahsen, H. (2000). Brain potentials in the processing of complex sentences: an ERP study of control and raising constructions. *Journal of Psycholinguistic Research*, *29*(2), 141–154.

Flege, J. E., & Liu, S. (2001). The effect of experience on adults' acquisition of a second language. *Studies in Second Language Acquisition*, *23*(4), 527–552.

Fox, J., & Weisberg, S. (2011). *An {R} Companion to Applied Regression, Second Edition*. Thousand Oaks, CA: Sage.

Frank, S. L., Otten, L. J., Galli, G., & Vigliocco, G. (2015). The ERP response to the amount of information conveyed by words in sentences. *Brain and Language*, *140*, 1–11. DOI:10.1016/j.bandl.2014.10.006

Frazier, L. (1987). Sentence processing: a tutorial review. In M. Coltheart (Ed.), *Attention and Performance XII: The Psychology of Reading* (pp. 559–586). Hillsdale, NJ: Lawrence Erlbaum.

Frenck-Mestre, C. (2002). An on-line look at sentence processing in the second language. In R. Heredia & J. Altarriba (Eds.), *Bilingual sentence processing* (pp. 217–236). Elsevier.

Friederici, A. D. (1995). The time course of syntactic activation during language processing: A model based on neuropsychological and neurophysiological data. *Brain and Language*, *50*, 259–281. DOI:10.1006/brln.1995.1048

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *TRENDS in Cognitive Sciences*, *6*(2), 78–84. DOI:10.1016/S1364-6613(00)01839-8

Friederici, A. D. (2011). The brain basis of language processing: from structure to function. *Physiological Reviews*, *91*(4), 1357–1392. DOI:10.1152/physrev.00006.2011

Friederici, A. D. (2012). The cortical language circuit: from auditory perception to sentence comprehension. *Trends in Cognitive Sciences*, *16*(5), 262–268. DOI:10.1016/j.tics.2012.04.001

Friederici, A. D., Gunter, T. C., Hahne, A., & Mauth, K. (2004a). The relative timing of syntactic and semantic processes in sentence comprehension. *NeuroReport*, *15*(1), 165–169. DOI:DOI:10.1097/00001756-200401190-00032

Friederici, A. D., Gunter, T. C., Hahne, A., & Mauth, K. (2004b). The relative timing of syntactic and semantic processes in sentence comprehension. *NeuroReport*, *15*(1), 165–169. DOI:10.1097/01.wnr.0000093585.33576.65

Friederici, A. D., Gunter, T. C., Hahne, A., & Mauth, K. (2004c). The Relative Timing of Syntactic and Semantic Processes in sentences. *NeuroReport*, *15*(1), 165–169. DOI:10.1097/01.wnr.0000093585.33576.65

Friederici, A. D., Hahne, A., & Mecklinger, A. (1996). Temporal structure of syntactic parsing: Early and late event-related brain potential effects. *Journal of Experimental Psychology: Learning Memory and Cognition*, *22*(5), 1219–1248. DOI:10.1037/0278-7393.22.5.1219

Friederici, A. D., & Kotz, S. A. (2003). The brain basis of syntactic processes: functional imaging and lesion studies. *NeuroImage*, *20*(SUPPL. 1), 8–17. DOI:10.1016/j.neuroimage.2003.09.003

Friederici, A. D., Mecklinger, A., Spencer, K. M., Steinhauer, K., & Donchin, E. (2001). Syntactic parsing preferences and their on-line revisions: A spatio-temporal analysis of event-related brain potentials. *Cognitive Brain Research*, *11*, 305–323. DOI:10.1016/S0926-6410(00)00065-3

Friederici, A. D., Pfeifer, E., & Hahne, A. (1993). Event-related brain potentials during natural speech processing: effects of semantic, morphological and syntactic violations. *Cognitive Brain Research*, *1*(3), 183–192. DOI:10.1016/0926-6410(93)90026-2

Friederici, A. D., & Singer, W. (2015). Grounding language processing on basic neurophysiological principles. *Trends in Cognitive Sciences*, *19*(6), 329–338. DOI:10.1016/j.tics.2015.03.012

Friederici, A. D., Steinhauer, K., & Frisch, S. (1999). Lexical integration: sequential effects of syntactic and semantic information. *Memory and Cognition*, *27*(3), 438–453. DOI:10.3758/BF03211539

Friederici, A. D., Steinhauer, K., & Pfeifer, E. (2002). Brain signatures of artificial language processing: evidence challenging the critical period hypothesis, *99*(1), 529–534.

Frisch, S., Hahne, A., & Friederici, A. D. (2004). Word category and verb-argument structure information in the dynamics of parsing. *Cognition*, *91*, 191–219. DOI:10.1016/j.cognition.2003.09.009

Gelman, A., & Hill, J. (2007). *Data Analysis Using Regression and Multilever/Hierarchical Models*. Cambridge University Press.

Goucha, T., Zaccarella, E., & Friederici, A. D. (2017). A revival of Homo loquens as a builder of labeled structures: neurocognitive considerations. *Neuroscience and Biobehavioral Reviews*, *81*, 213–224. DOI:10.1016/j.neubiorev.2017.01.036

Greiffenstein, M. F., Baker, W. J., & Gola, T. (1994). Validation of malingered amnesia measures with a large clinical sample. *Psychological Assessment*, *6*(3), 218–224. DOI:10.1037/1040-3590.6.3.218

Green, D. W. (2003). Neural basis of lexicon and grammar in L2 acquisition. In R. van Hout, A. Hulk, F. Kuiken, & T. R.J. (Eds.), *The Lexicon–Syntax Interface in Second Language Acquisition* (pp. 197–218). Amsterdam, the Netherlands: Benjamins.

Guajardo, L. F., & Wicha, N. Y. Y. (2014). Morphosyntax can modulate the N400 component : event related potentials to gender-marked post-nominal adjectives. *NeuroImage*, *91*, 262–272. DOI:10.1016/j.neuroimage.2013.09.077

Gunter, T. C., Friederici, A. D., & Schriefers, H. (2000). Syntactic gender and semantic expectancy: ERPs reveal early autonomy and late interaction. *Journal of Cognitive Neuroscience*, *12*(4), 556–568. DOI:10.1162/089892900562336

Hagoort, P. (2003a). How the brain solves the binding problem for language: a neurocomputational model of syntactic processing. *NeuroImage*, *20*(SUPPL. 1), 18–29. DOI:10.1016/j.neuroimage.2003.09.013

Hagoort, P. (2003b). Interplay between syntax and semantics during sentence comprehension: ERP effects of combining syntactic and semantic violations. *Journal of Cognitive Neuroscience*, *15*(6), 883–899. DOI:10.1162/089892903322370807

Hagoort, P. (2005). On Broca, brain, and binding: a new framework. *Trends in Cognitive Sciences*, *9*(9), 416–423. DOI:10.1016/j.tics.2005.07.004

Hagoort, P. (2016). The memory, unification, and control (MUC) model of language. In G. Hickok & S. Small (Eds.), *Neurobiology of Language* (pp. 339–347). Elsevier Inc. DOI:10.4324/9780203968512

Hagoort, P., Baggio, G., & Willems, R. M. (2009). Semantic Unification. In M. S. Gazzaniga (Ed.), *The Cognitive Neurosciences* (4th ed., pp. 1–13). Cambridge, MA: MIT Press. DOI:10.1007/978-3-642-54789-8_1

Hagoort, P., Brown, C., & Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language and Cognitive Processes*, *8*(4), 439–483. DOI:10.1080/01690969308407585

Hagoort, P., Brown, C. M., & Osterhout, L. (1999). The neurocognition of syntactic processing. In C. M. Brown & P. Hagoort (Eds.), *The neurocognition of language* (pp. 273–317). Oxford University Press.

Hagoort, P., Hald, L., Bastiaansen, M., & Petersson, K. M. (2004). Integration of word meaning and world knowledge in language comprehension. *Science*, *304*, 438–441. Retrieved from http://science.sciencemag.org/content/304/5669/438

Hagoort, P., Wassenaar, M., & Brown, C. M. (2003). Syntax-related ERP-effects in Dutch. *Cognitive Brain Research*, *16*, 38–50.

Hagoort, Peter. (2003). Interplay between syntax and semantics during sentence comprehension: ERP effects of combining syntactic and semantic violations. *Journal of Cognitive Neuroscience*, *15*(6), 883–899. DOI:10.1162/089892903322370807

Hahne, A. (2001). What's different in second-language processing? Evidence from event-related potentials. *Journal of Psycholinguistic Research*, *30*(3), 251–266. DOI:10.1023/A:1010490917575

Hahne, A., & Friederici, A. D. (1999). Electrophysiological evidence for two steps in syntactic analysis: early automatic and late controlled processes. *Journal of Cognitive Neuroscience*, *11*(2), 194–205.

Hahne, A., & Friederici, A. D. (2001). Processing a second language: late learners' comprehension mechanisms as revealed by event-related brain potentials. *Bilingualism: Language and Cognition*, *4*(02). DOI:10.1017/S1366728901000232

Hahne, A., & Friederici, A. D. (2002). Differential task effects on semantic and syntactic processes as revealed by ERPs. *Cognitive Brain Research*, *13*(3), 339–356.

Harrell, F. E. J. (2019). Package 'Hmisc': version 4.2-0. Retrieved from https://github.com/harrelfe/Hmisc

Hartshorne, J. K., Tenenbaum, J. B., & Pinker, S. (2018). A critical period for second language acquisition: evidence from 2/3 million English speakers. *Cognition*, *177*(May), 263–277. DOI:10.1016/j.cognition.2018.04.007

Hasting, A. S., & Kotz, S. A. (2008). Speeding up syntax: on the relative timing and automaticity of local phrase structure and morphosyntactic processing as reflected in event-related brain potentials. *Journal of Cognitive Neuroscience*, *20*(7), 1207–1219. DOI:10.1162/jocn.2008.20083

Hernandez, A., Li, P., & MacWhinney, B. (2005). The emergence of competing modules in bilingualism. *Trends in Cognitive Sciences*, *9*(5), 220–225. DOI:doi:10.1016/j.tics.2005.03.003

Hinojosa, J. A., Martin-Loeches, M., Casado, P., Muñoz, F., & Rubia, F. J. (2003). Similarities and differences between phrase structure and morphosyntactic violations in Spanish: an event-related potentials study. *Language and Cognitive Processes*, *18*(2), 113–142. DOI:10.1080/01690960143000489

Hopp, H. (2016). The timing of lexical and syntactic processes in L2 sentence comprehension. *Applied Psycholinguistcs*, 37, 1253–1280.

Hothorn, T., Hornik, K., & Zeileis, A. (2006). Unbiased Recursive Partitioning: A Conditional Inference Framework. *Journal of Computational and Graphical Statistics*, *15*(3), 651–674.

Hothorn, T., & Zeileis, A. (2015). partykit: A modular toolkit for recursive partytioning in R. *Journal of Machine Learning Research*, *16*, 3905–3909.

Isel, F. (2007). Syntactic and referential processes in second-language learners: Event-related brain potential evidence. *NeuroReport*, *18*(18), 1885–1889. DOI:10.1097/WNR.0b013e3282f1d518

Isel, F., Hahne, A., Maess, B., & Friederici, A. D. (2007). Neurodynamics of sentence interpretation: ERP evidence from French. *Biological Psychology*, *74*(3), 337–346. DOI:10.1016/j.biopsycho.2006.09.003

Jackendoff, R. S. (2002). *Foundations of language: Brain, meaning, grammar, evolution*. Oxford University Press.

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: the influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, *21*, 60–99. DOI:10.1016/0010-0285(89)90003-0

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: individual differences in working memory. *Psychological Review*, *99*, 122–149.

Kaan, E., Harris, A., Gibson, E., & Holcomb, P. (2000). The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes*, *15*(2), 159–201.

Kaan, E., & Swaab, T. Y. (2003). Repair, revision, and complexity in syntactic analysis : an electrophysiological differentiation. *Journal of Cognitive Neuroscience*, *15*(1), 98–110.

Kasparian, K., Vespignani, F., & Steinhauer, K. (2017). First language attrition induces changes in online morphosyntactic processing and re-analysis : an ERP study of number agreement in complex italian sentences. *Cognitive Science*, *41*, 1760–1803. DOI:10.1111/cogs.12450

Kiefer, M. (2002). The N400 is modulated by unconsciously perceived masked words: further evidence for an automatic spreading activation account of N400 priming effects. *Cognitive Brain Research*, *13*, 27–39. DOI:10.1016/s0926-6410(01)00085-4

Kim, A., Oines, L., & Miyake, A. (2018). Individual differences in verbal working memory underlie a tradeoff between semantic and structural processing difficulty during language comprehension: an ERP investigation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(3), 406–420. DOI:10.1037/xlm0000457

Kim, A., Oines, L., & Sikos, L. (2016). Prediction during sentence comprehension is more than a sum of lexical associations: the role of event knowledge. *Language, Cognition and Neuroscience*, *31*(5), 597–601. DOI:10.1080/23273798.2015.1102950

Kim, A., & Osterhout, L. (2005). The independence of combinatory semantic processing: evidence from event-related potentials. *Journal of Memory and Language*, *52*(2), 205–225. DOI:10.1016/j.jml.2004.10.002

Kimball, J. (1975). Predictive analysis and over-the-top parsing. In J. Kimball (Ed.), *Syntax and Semantics* (pp. 155–179). New York: Academic Press.

Kos, M., van den Brink, D., & Hagoort, P. (2012). Individual variation in the late positive complex to semantic anomalies. *Frontiers in Psychology*, *3*, 318. DOI:10.3389/fpsyg.2012.00318

Kuhn, M. (2008). Building predictive models in R using the caret package. *Journal of Statistical Software*, *28*(5), 1–26.

Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: Challenges to syntax. *Brain Research*, *1146*(1), 23–49. DOI:10.1016/j.brainres.2006.12.063

Kuperberg, G. R., Caplan, D., Sitnikova, T., Eddy, M., & Holcomb, P. J. (2006). Neural correlates of processing syntactic, semantic, and thematic relationships in sentences. *Language and Cognitive Processes*, *21*(5), 489–530. DOI:10.1080/01690960500094279

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: finding meaning in the N400 component of the event related brain potential (ERP). *Annual Review of Psychology*, *62*(1), 621–647. DOI:10.1146/annurev.psych.093008.131123

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: brain potentials reflect semantic incongruity. *Science*, *207*(4427), 203–205.

Kutas, M., & Hillyard, S. A. (1983). Event-related brain potentials to grammatical errors and semantic anomalies. *Memory & Cognition*, *11*(5), 539–550. DOI:10.3758/BF03196991

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26. DOI:doi: 10.18637/jss.v082.i13

Labelle, M. (2001). Trente ans de psycholinguistique. *Revue Québécoise de Linguistique*, *30*(1), 155–176.

Lago, S., Sloggett, S., Schlueter, Z., Chow, W. Y., Williams, A., Lau, E., & Phillips, C. (2017). Coreference and Antecedent Representation Across Languages. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(5), 795-817. DOI:10.1037/xlm0000343

Lau, E. F., Holcomb, P. J., & Kuperberg, G. R. (2013). Dissociating N400 effects of prediction from association in single word contexts. *Journal of Cognitive Neuroscience*, *25*(3), 484–502. DOI:doi:10.1162/jocn_a_00328

Lau, E., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (de)constructing the N400. *Nature Reviews Neuroscience*, *9*(12), 920–933. DOI:10.1038/nrn2532

Lau, E., Stroud, C., Plesch, S., & Phillips, C. (2006). The role of structural prediction in rapid syntactic analysis. *Brain and Language*, *98*(1), 74–88. DOI:10.1016/j.bandl.2006.02.003

Leclercq, P., Edmonds, A., & Hilton, H. (2014). *Measuring L2 Proficiency: Perspectives from SLA.* Clevedon: Multilingual Matters Ltd.

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: a quick and valid lexical test for advanced learners of English. *Behavior Research Methods*, *44*(2), 325–343. DOI:10.3758/s13428-011-0146-0

Leminen, A., Smolka, E., Duñabeitia, J. A., & Pliatsikas, C. (2018). Morphological processing in the brain: the good (inflection), the bad (derivation) and the ugly (compounding). *Cortex*, (September). DOI:10.1016/j.cortex.2018.08.016

Lenneberg, E. H. (1967). *Biological foundations of language* (Wiley). New York.

Lenth, R. (2018). emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.1.3. Retrieved from https://cran.r-project.org/package=emmeans

Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R News*, *2*(3), 18–22.

Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: An open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience*, *8*, 213.

Luck, S. (2005). *An introduction to the event-related potential technique*. MIT Press.

Luo, Y., Zhang, Y., Feng, X., & Zhou, X. (2010). Electroencephalogram oscillations differentiate semantic and prosodic processes during sentence reading. *Neuroscience*, *169*(2), 654–664. DOI:10.1016/j.neuroscience.2010.05.032

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). Lexical nature of syntactic ambiguity resolution. *Psychological Review*, *101*(4), 676–703.

MacLennan, H. (1945). *Two Solitudes*. Toronto: Collins.

MacWhinney, B. (2005). A unified model of language acquisition. In *Handbook of bilingualism: Psycholinguistic approaches* (p. 4967).

Mancini, S., Molinaro, N., Rizzi, L., & Carreiras, M. (2011). When persons disagree: an ERP study of Unagreement in Spanish. *Psychophysiology*, *48*(10), 1361–1371. DOI:10.1111/j.1469-8986.2011.01212.x

Marquis, A., & Royle, P. (2019). Verb acquisition in monolingual and multilingual children and adults. In P. Guijarro-Fuentes & C. Suárez-Gómez (Eds.), *Proceedings of GALA 2017: Language Acquisition and Development* (pp. 307–324). Newcastle upon Tyne, UK: Cambridge Scholars Publishing.

Martín-Loeches, M., Nigbur, R., Casado, P., Hohlfeld, A., & Sommer, W. (2006). Semantics prevalence over syntax during sentence processing : a brain potential study of noun – adjective agreement in Spanish. *Brain Research*, *3*, 178–189. DOI:10.1016/j.brainres.2006.03.094

Matsuki, K., Kuperman, V., & Van Dyke, J. A. (2016). The Random Forests statistical technique: an examination of its value for the study of reading. *Scientific Studies of Reading*, *20*(1), 20–33. DOI:10.1080/10888438.2015.1107073.The

Mecklinger, A., Schriefers, H., Steinhauer, K., & Friederici, A. D. (1995). Processing relative clauses varying on syntactic and semantic dimensions: an analysis with event-related potentials. *Memory & Cognition*, *23*(4), 477–494.

Meerendonk, N. Van De, Kolk, H. H. J., Vissers, C. T. W. M., & Chwilla, D. J. (2008). Monitoring in language perception : mild and strong conflicts elicit different ERP patterns. *Journal of Cognitive Neuroscience*, *22*(1), 67–82.

Mehravari, A. S., Tanner, D., Wampler, E. K., Valentine, G. D., & Osterhout, L. (2015). Effects of grammaticality and morphological complexity on the P600 event-related potential component. *PLoS ONE*, *10*(10), 1–16. DOI:10.1371/journal.pone.0140850

Meulman, N., Wieling, M., Sprenger, S. A., Stowe, L. A., & Schmid, M. S. (2015). Age effects in L2 grammar processing as revealed by ERPs and how (not) to study them. *Plos One*, *10*(12), e0143328. DOI:10.1371/journal.pone.0143328

Molinaro, N., Barber, H. A., Caffarra, S., & Carreiras, M. (2015). On the left anterior negativity (LAN): the case of morphosyntactic agreement: a reply to Tanner et al. *Cortex*, *66*, 156–159. DOI:10.1016/j.cortex.2014.06.009

Molinaro, N., Barber, H. A., & Carreiras, M. (2011). Grammatical agreement processing in reading: ERP findings and future directions. *Cortex*, *47*(8), 908–930. DOI:10.1016/j.cortex.2011.02.019

Molinaro, N., Giannelli, F., Caffarra, S., & Martin, C. (2017). Hierarchical levels of representation in language prediction: The influence of first language acquisition in highly proficient bilinguals. *Cognition*, *164*, 61–73. DOI:10.1016/j.cognition.2017.03.012

Molinaro, N., Monsalve, I. F., & Lizarazu, M. (2016). Is there a common oscillatory brain mechanism for producing and predicting language? *Language, Cognition and*

Neuroscience, *31*(1), 145–158. DOI:10.1080/23273798.2015.1077978

Molinaro, N., Vespignani, F., Zamparelli, R., & Job, R. (2011). Why brother and sister are not just siblings: Repair processes in agreement computation. *Journal of Memory and Language*, *64*(3), 211–232. DOI:10.1016/j.jml.2010.12.002

Morgan-Short, K., Steinhauer, K., Sanz, C., & Ullman, M. T. (2012). Explicit and implicit second language training differentially affect the achievement of native-like brain activation patterns. *Jordan Journal of Civil Engineering*, *24*(4), 933–947. DOI:10.1162/jocn_a_00119

Mueller, J. L. (2005). Electrophysiological correlates of second language processing. *Second Language Research*, *21*(2), 152–174. DOI:10.1191/0267658305sr256oa

Mueller, J. L., Hahne, A., Fujii, Y., & Friederici, A. D. (2005). Native and nonnative speakers' processing of a miniature version of Japanese as revealed by ERPs. *Journal of Cognitive Neuroscience*, *17*(8), 1229–1244. DOI:10.1162/0898929055002463

Neville, H., Nicol, J. L., Barss, A., Forster, K. I., & Garrett, M. F. (1991). Syntactically based sentence processing classes: Evidence from event-related brain potentials. *Journal of Cognitive Neuroscience*, *3*(2), 151–165. DOI:10.1162/jocn.1991.3.2.151

Newman, A. J., Ullman, M. T., Pancheva, R., Waligura, D. L., & Neville, H. J. (2007). An ERP study of regular and irregular English past tense inflection. *NeuroImage*, *34*, 435–445. DOI:10.1016/j.neuroimage.2006.09.007

Newport, E. L., Bavelier, D., & Neville, H. J. (2001). Perspectives on a Critical Period for Language Acquisition. In E. Dupoux (Ed.), *Language, brain, and cognitive development* (pp. 481–502). MIT Press. DOI:10.1067/mhn.2001.115372

Nichols, E. S., & Joanisse, M. F. (2019). Individual differences predict ERP signatures of second language learning of novel grammatical rules. *Bilingualism: Language and Cognition*, *22*(1), 78–92. DOI:10.1017/S1366728917000566

Nickels, S. (2016). *A critical re-evaluation of Friederici's (2002) sentence processing model with particular consideration of task effects, item effects, and inter-individual differences*. McGill University, Montreal.

Nickels, S., Bokhari, F. S., & Steinhauer, K. (2014). "semantic blocking"solved at last: ERP evidence challenging syntax-first models. In *Annual Meeting of the Society for the Neurobiology of Language*.

Nickels, S., & Steinhauer, K. (2018). Prosody–syntax integration in a second language: contrasting event-related potentials from German and Chinese learners of English using linear mixed effect models. *Second Language Research*, *34*(1), 9–37. DOI:10.1177/0267658316649998

Nieuwland, M. S., & Kuperberg, G. R. (2008). When the truth is not too hard to handle: an ERP study on the pragmatics of negation. *Psychological Science*, *19*(12), 1213–1218. DOI:10.1111/j.1467-9280.2008.02226.x

Nieuwland, M. S., Martin, A. E., & Carreiras, M. (2013). Event-related brain potential evidence for animacy processing asymmetries during sentence comprehension. *Brain and*

Language, *126*(2), 151–158. DOI:10.1016/j.bandl.2013.04.005

Nieuwland, M. S., Politzer-Ahles, S., Heyselaar, E., Segaert, K., Darley, E., Kazanina, N., … Huettig, F. (2018). Large-scale replication study reveals a limit on probabilistic prediction in language comprehension. *ELife*, *7*, 1–24. DOI:10.7554/elife.33468

Ojima, S., Nakata, H., & Kakigi, R. (2005). An ERP study of second language learning after childhood: effects of proficiency. *Journal of Cognitive Neuroscience*, *17*(8), 1212–1228.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburg inventory. *Neuropsychologia*, *9*, 97–113. DOI:10.1016/0028-3932(71)90067-4

Osterhout, L. (1997). On the brain response to syntactic anomalies: Manipulations of word position and word class reveal individual differences. *Brain and Language*, *59*(3), 494–522. DOI:10.1006/brln.1997.1793

Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, *36*(6), 785–806.

Osterhout, L., Holcomb, P. J., & Swinney, D. A. (1994). Brain potentials elicited by garden-path sentences: evidence of the application of verb information during parsing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(4), 786–803.

Osterhout, L., McLaughlin, J., Kim, A., Greenwald, R., & Inoue, K. (2004). Sentences in the brain: event-related potentials as real-time reflections of sentence comprehension and language learning. In M. Carreiras & C. Clifton (Eds.), *The On-line Study of Sentence Comprehension: Eye-Tracking, ERPs and Beyond* (pp. 271–308). Psychology Press.

Osterhout, L., Mclaughlin, J., Pitk, I., Frenck-mestre, C., & Molinaro, N. (2006). Novice learners, longitudinal designs , and event-related potentials: a means for exploring the neurocognition of second language processing. *Language Learning*, *56*(s1), 199–230. DOI:10.1111/j.1467-9922.2006.00361.x

Osterhout, L., & Mobley, L. A. (1995). Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language*. DOI:10.1006/jmla.1995.1033

Otten, M., Nieuwland, M. S., & Van Berkum, J. J. A. (2007). Great expectations: Specific lexical anticipation influences the processing of spoken language. *BMC Neuroscience*, *8*, 1–9. DOI:10.1186/1471-2202-8-89

Pakulak, E., & Neville, H. J. (2011). Maturational Constraints on the Recruitment of Early Processes for Syntactic Processing. *Journal of Cognitive Neuroscience*, *23*(10), 2752–2765. DOI:10.1162/jocn.2010.21586.Maturational

Perception Research Systems. (2007). Paradigm Stimulus Presentation.

Phillips, C., Kazanina, N., & Abada, S. H. (2005). ERP effects of the processing of syntactic long-distance dependencies. *Cognitive Brain Research*, *22*(3), 407–428. DOI:10.1016/j.cogbrainres.2004.09.012

Piai, V., Meyer, L., Schreuder, R., & Bastiaansen, M. C. M. (2013). Sit down and read on: working memory and long-term memory in particle-verb processing. *Brain and Language*, *127*(2), 296–306. DOI:10.1016/j.bandl.2013.09.015

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, *118*(10), 2128–2148. DOI:10.1016/j.clinph.2007.04.019.Updating

Pollard, C., & Sag, I. A. (1994). *Head-Driven Phrase Structure Grammar*. Chicago, IL: University of Chicago Press.

Poschmann, C., & Wagner, M. (2016). Relative clause extraposition and prosody in German. *Natural Language and Linguistic Theory*, *34*(3), 1021–1066. DOI:10.1007/s11049-015-9314-8

Qi, Z., Beach, S. D., Finn, A. S., Minas, J., Goetz, C., Chan, B., & Gabrieli, J. D. E. (2017). Native-language N400 and P600 predict dissociable language-learning abilities in adults. *Neuropsychologia*, *98*, 177–191. DOI:10.1016/j.neuropsychologia.2016.10.005

R Core Team. (2017). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Robert, C., & Rico Duarte, L. (2016). Semantic richness and aging: the effect of number of features in the lexical decision task. *Journal of Psycholinguistic Research*, *45*(2), 359–365. DOI:10.1007/s10936-015-9352-8

Rossi, S., Gugler, M. F., Friederici, A. D., & Hahne, A. (2006). The impact of proficiency on syntactic second-language processing of German and Italian: evidence from event-related potentials. *Journal of Cognitive Neuroscience*, *18*(12), 2030–2048. DOI:10.1162/jocn.2006.18.12.2030

Rossi, S., Gugler, M. F., Hahne, A., & Friederici, A. D. (2005). When word category information encounters morphosyntax: an ERP study. *Neuroscience Letters*, *384*, 228–233. DOI:10.1016/j.neulet.2005.04.077

Royle, P., & Courteau, E. (2014). Language processing in children with specific language impairment: a review of event-related potential studies. In L. T. Klein & V. Amato (Eds.), *Language Processing: New Research* (pp. 33–64). Hauppauge, NY: Nova Science Publishers.

Royle, P., Drury, J. E., Bourguignon, N., & Steinhauer, K. (2012). The temporal dynamics of inflected word recognition: a masked ERP priming study of French verbs. *Neuropsychologia*, *50*(14), 3542–3553. DOI:10.1016/j.neuropsychologia.2012.09.007

Royle, P., Drury, J. E., & Steinhauer, K. (2013). ERPs and task effects in the auditory processing of gender agreement and semantics in French. In *The Mental Lexicon* (Vol. 8, pp. 216–244). DOI:doi:10.1075/ml.8.2.05roy

Royle P., Fromont L.A., Drury J.E. (2018). Definiteness and maximality in French language acquisition, more adult-like than you would expect. *Frontiers in Communication*, 3, 27. DOI: 10.3389/fcomm.2018.00027

Royle, P., Steinhauer, K., Dessureault, É., Herbay, A. C., & Brambati, S. M. (2019). Aging and language: maintenance of morphological representations in older adults. *Frontiers in Communication*, *4*, 16. DOI:10.3389/fcomm.2019.00016

Sanders, L. D., & Neville, H. J. (2003). An ERP study of continuous speech processing II. Segmentation, semantics, and syntax in native speakers. *Cognitive Brain Research*, *15*, 228–240. DOI:10.1016/S0926-6410(02)00195-7

Sassenhagen, J., Schlesewsky, M., & Bornkessel-Schlesewsky, I. (2014). The P600-as-P3 hypothesis revisited: single-trial analyses reveal that the late EEG positivity following linguistically deviant material is reaction time aligned. *Brain and Language*, *137*, 29–39. DOI:10.1016/j.bandl.2014.07.010

Sereno, S. C., & Rayner, K. (2003). Measuring word recognition in reading: eye movements and event-related potentials. *Trends in Cognitive Sciences*, *7*(11), 489–493. DOI:10.1016/j.tics.2003.09.010

Silva-Pereyra, J., Rivera-Gaxiola, M., & Kuhl, P. K. (2005). An event-related brain potential study of sentence comprehension in preschoolers: semantic and morphosyntactic processing. *Cognitive Brain Research*, *23*, 247–258. DOI:10.1016/j.cogbrainres.2004.10.015

Silver, N. C., Hittner, J. B., & May, K. (2004). Testing dependent correlations with nonoverlapping variables: a monte carlo simulation. *Journal of Experimental Education*, *73*, 53–69. DOI:10.3200/JEXE.71.1.53-70

Soylu, F. (2010). Forward/backward digit-span task. *Archives of Neurobehavioral Experiments and Stimuli*, *218*. Retrieved from http://www.neurobs.com/ex_files/expt_view?id=218

Squires, N., Squires, K., & Hillyard, S. (1975). Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man. *Electroencephalogr Clin Neurophysiol*, *38*(4), 387–401.

Steinhauer, K. (2014). Event-related potentials (ERPs) in second language research: a brief introduction to the technique, a selected review, and an invitation to reconsider critical periods in L2. *Applied Linguistics*, *35*(4), 393–417. DOI:10.1093/applin/amu028

Steinhauer, K., & Drury, J. E. (2012). On the early left-anterior negativity (ELAN) in syntax studies. *Brain and Language*, *120*(2), 135–162. DOI:10.1016/j.bandl.2011.07.001

Steinhauer, K., Drury, J. E., Portner, P., Walenski, M., & Ullman, M. T. (2010). Syntax, concepts, and logic in the temporal dynamics of language comprehension: evidence from event-related potentials. *Neuropsychologia*, *48*(6), 1525–1542. DOI:10.1016/j.neuropsychologia.2010.01.013

Steinhauer, K., Royle, P., Drury, J. E. J. E., & Fromont, L. A. (2017). The priming of priming : evidence that the N400 reflects context-dependent post-retrieval word integration in working memory. *Neuroscience Letters*, *651*, 192–197. DOI:10.1016/j.neulet.2017.05.007

Steinhauer, K., White, E. J., & Drury, J. E. (2009). Temporal dynamics of late second language acquisition: evidence from event-related brain potentials. *Second Language Research*, *25*(1), 13–41. DOI:10.1177/0267658308098995

Stowe, L. A., Kaan, E., Sabourin, L., & Taylor, R. C. (2018). The sentence wrap-up dogma. *Cognition*, *176*(March), 232–247. DOI:10.1016/j.cognition.2018.03.011

Strobl, C., Boulesteix, A., Kneib, T., Augustin, T., & Zeileis, A. (2008). Conditional variable importance for random forests, *11*, 1–11. DOI:10.1186/1471-2105-9-307

Strobl, C., Malley, J., & Tutz, G. (2009). An introduction to recursive partitioning : rationale , application, and characteristics of classification and regression trees, bagging, and Random Forests, *14*(4), 323–348. DOI:10.1037/a0016973

Tagliamonte, S. A., & Baayen, R. H. (2012). Models, forests, and trees of York English: Was/were variation as a case study for statistical practice. *Language Variation and Change*, *24*, 135–178. DOI:10.1017/S0954394512000129

Tanner, D. (2015). On the left anterior negativity (LAN) in electrophysiological studies of morphosyntactic agreement: a commentary on "Grammatical agreement processing in reading: ERP findings and future directions" by Molinaro et al., 2014. *Cortex*, *66*, 149–155. DOI:10.1016/j.cortex.2014.04.007

Tanner, D. (2019). Robust neurocognitive individual differences in grammatical agreement processing: a latent variable approach. *Cortex*, *111*, 210–237. DOI:10.1016/j.cortex.2018.10.011

Tanner, D., Inoue, K., & Osterhout, L. (2014). Brain-based individual differences in online L2 grammatical comprehension. *Bilingualism: Language and Cognition*, *17*(2), 277–293. DOI:10.1017/S1366728913000370

Tanner, D., McLaughlin, J., Herschensohn, J., & Osterhout, L. (2013). Individual differences reveal stages of L2 grammatical acquisition: ERP evidence. *Bilingualism: Language and Cognition*, *16*(2), 367–382. DOI:10.1017/S1366728912000302

Tanner, D., & Van Hell, J. G. (2014). ERPs reveal individual differences in morphosyntactic processing. *Neuropsychologia*, *56*(1), 289–301. DOI:10.1016/j.neuropsychologia.2014.02.002

Tomaschek, F., Hendrix, P., & Baayen, R. H. (2018). Strategies for addressing collinearity in multivariate linguistic data. *Journal of Phonetics*, *71*, 249–267. DOI:10.1016/j.wocn.2018.09.004

Tremblay, A., & Garrison, M. D. (2010). Cloze Tests : A Tool for Proficiency Assessment in Research on L2 French. In M. T. Prior, Y. Watanabe, & S.-K. Lee (Eds.), *Selected Proceedings of the 2008 Second Language Research Forum* (pp. 73–88). Somerville, MA, USA: Cascadilla Proceedings Project.

Tuijl, A. (2017). *Word categorization in Mandarin: with special attention to its system of classifiers and measure words*.

Ullman, M. T. (2001). A Neurocognitive Perspective on Language: The Declarative/Procedural Model. *Nature Reviews Neuroscience*, *2*(10), 717–726. DOI:10.1038/35094573

Ullman, M. T. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition*, *92*(1–2), 231–270. DOI:10.1016/j.cognition.2003.10.008

Valois, D. & Royle, P. (2009). Partitivity, atomization, and noun-drop: A longitudinal study of French child language. *Language Acquisition: A Journal of Developmental Linguistics,* *16*(2), 82-105. DOI: 10.1080/10489220902773079

Van Berkum, J. J. A., Brown, C. M., Hagoort, P., & Zwitserlood, P. (2003). Event-related brain potentials reflect discourse-referential ambiguity in spoken language comprehension. *Psychophysiology*, *40*(2), 235–248. DOI:10.1111/1469-8986.00025

Van den Brink, D., & Hagoort, P. (2004). The influence of semantic and syntactic context constraints on lexical selection and integration in spoken-word comprehension as revealed by ERPs. *Journal of Cognitive Neuroscience*, *16*(6), 1068–1084. DOI:10.1162/0898929041502670

Van Hell, J. G., & Abdollahi, F. (2017). Individual variation in syntactic processing in the second language: Electrophysiological approaches. *Developmental Perspectives in Written Language and Literacy*, (February 2018), 257–273. DOI:10.1075/z.206.16van

Van Hell, J. G., & Tanner, D. (2012). Second language proficiency and cross-language lexical activation. *Language Learning*, *62*(s2), 148–171. DOI:10.1111/j.1467-9922.2012.00710.x

Van Petten, C., & Kutas, M. (1990). Interactions between sentence context and word frequency in event-related brain potentials. *Memory & Cognition*, *18*(4), 380–393. DOI:10.3758/BF03197127

Van Petten, C., & Luka, B. J. (2006). Neural localization of semantic context effects in electromagnetic and hemodynamic studies. *Brain and Language*, *97*(3), 279–293. DOI:10.1016/j.bandl.2005.11.003

Vanhove, J. (2013). The critical period hypothesis in second language acquisition: a statistical critique and a reanalysis. *PLoS ONE*, *8*(7). DOI:10.1371/journal.pone.0069172

Vissers, C., Kolk, H. H. J., van de Meerendonk, N., & Chwilla, D. J. (2008). Monitoring in language perception: evidence from ERPs in a picture-sentence matching task. *Neuropsychologia*, *46*(4), 967–982. DOI:10.1016/j.neuropsychologia.2007.11.027

Vos, S. H., Gunter, T. C., Kolk, H. H. J., & Mulder, G. (2001). Working memory constraints on syntactic processing: An electrophysiological investigation. *Psychophysiology*, *38*(1), 41–63. DOI:10.1017/S0048577201980278

Vosse, T., & Kempen, G. (2000). Syntactic structure assembly in human parsing: a computational model based on competitive inhibition and a lexicalist grammar. *Cognition*, *75*, 105–143. DOI:10.1016/S0010-0277(00)00063-9

Weber-Fox, C. M., & Neville, H. J. (1996). Maturational constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. *Journal of Cognitive Neuroscience*, *8*(3), 231–256.

White, E. J., Genesee, F., White, L., King, E., & Steinhauer, K. (2006). Factors contributing to native-like processing of a late-acquired second language: evidence from event-related brain potentials. In *41st Annual Boston University Conference on Language Development*.

White, L. (2003). The initial state. In *Second Language Acquisition and Universal Grammar* (pp. 58–99). Cambridge: Cambridge University Press.

DOI:10.1017/CBO9780511815065.004

Wright, B., & Garrett, M. (1984). Lexical decision in sentences: effects of syntactic structure. *Memory & Cognition*, *12*(1), 31–45.

Wright, M., & Ziegler, A. (2017). for High Dimensional Data in C ++ and R, *77*(1), 1–17. DOI:10.18637/jss.v077.i01

Yang, Y., Wu, F., & Zhou, X. (2015). Semantic processing persists despite anomalous syntactic category: ERP evidence from Chinese passive sentences. *PLoS ONE*, *10*(6), e0131936. DOI:10.1371/journal.pone.0131936

Zhang, Y., Li, P., Piao, Q., Liu, Y., Huang, Y., & Shu, H. (2013). Syntax does not necessarily precede semantics in sentence processing: ERP evidence from Chinese. *Brain and Language*, *126*(1), 8–19. DOI:10.1016/j.bandl.2013.04.001

Public repository

The present thesis, manuscripts, related studies, as well as all scripts and outputs are uploaded in the following repository: https://osf.io/k7vxp/.

Appendix 1 – Demographics and language background questionnaires

Profil du participant

*Required

1. **CODE** *

2. **Date de naissance (jour/mois/année)** *

3. **Sexe** *
   Mark only one oval.
   - F
   - M

4. **Occupation** *

5. **Accepteriez-vous que l'on vous recontacte pour d'autres études dans notre laboratoire de recherche ?** *
   Mark only one oval.
   - Oui
   - Non

6. **Avez-vous déjà eu des difficultés d'apprentissage ?** *
   Mark only one oval.
   - Oui
   - Non

7. **Avez-vous déjà passé un examen neurologique?** *
   Mark only one oval.
   - Oui
   - Non

8. **Si oui, expliquez pourquoi:**

9. **À quel âge avez-vous commencé l'école?** *

10. **Quel est votre dernier niveau de scolarité ?** *
Mark only one oval.

- Secondaire
- D.E.P.
- CÉGEP
- Université : Baccalauréat
- Maîtrise
- Doctorat

11. **À quelle fréquence vous lisait-on des histoires étant petit?** *
Mark only one oval.

	1	2	3	4	5	
jamais						tous les jours

12. **À quelle fréquence lisez-vous dans votre vie professionelle/étudiante?** *
(exemples: rapports, articles, livres, etc.)
Mark only one oval.

	1	2	3	4	5	
jamais						tous les jours

13. **À quelle fréquence lisez-vous dans votre vie personnelle?** *
(exemples: magazines, comiques, livres, etc.)
Mark only one oval.

	1	2	3	4	5	
jamais						tous les jours

14. **Combien de temps vous a-t-il fallu pour vous endormir la nuit dernière?** *
Mark only one oval.

- Très peu de temps
- Peu de temps
- Moyennement de temps
- Longtemps
- Très longtemps (je suis resté éveillé très longtemps)
- Ne sait pas

15. **Avez-vous bien dormi ?** *
Mark only one oval.

- Oui, de façon parfaite (d'un sommeil paisible, sans réveil nocturne)
- Oui, bien
- Moyennement bien
- Non, mal
- Non, très mal (sommeil agité, réveils fréquents)
- Ne sait pas

16. **Combien de temps avez-vous dormi ?** *
Mark only one oval.

- Très longtemps (je ne me suis pas réveillé spontanément)
- Longtemps
- Moyennement longtemps
- Peu de temps
- Très peu de temps (je me suis réveillé beaucoup trop tôt)
- Ne sait pas

17. **Vous êtes-vous réveillé au cours de la nuit ?** *
Mark only one oval.

- Jamais (j'ai dormi d'une seule traite)
- Rarement
- Relativement souvent
- Souvent
- Très souvent (réveils répétés)
- Ne sait pas

18. **Avez-vous fait des rêves ?** *
Mark only one oval.

- Aucun
- Quelques uns seulement
- Modérément
- Beaucoup
- Enormément et des rêves particulièrement marquants
- Ne sait pas

19. **Comment vous sentez-vous actuellement ?** *
Mark only one oval.

- En excellente forme
- En bonne forme
- Moyennement en forme
- En mauvaise forme
- En très mauvaise forme : fatigué, abattu
- Ne sait pas

Questionnaire – Utilisation des langues

Prénom et nom :____

Langue première de votre mère : ___________________ Âge d'acquisition_________

Langue seconde de votre mère : ___________________ Âge d'acquisition__________

Langue première de votre père : ___________________ Âge d'acquisition__________

Langue seconde de votre père : ___________________ Âge d'acquisition__________

Langues connues et comprises, de la mieux maîtrisée à la moins connue :

Indiquez, en pourcentage, votre exposition quotidienne à chacune de ces langues à chacune de ces périodes de votre vie.

Âge		**Français**	**Anglais**	**Autre : _____**	**Total**
0-5 ans	**École**				100%
	Maison				100%
5-11 ans	**École**				100%
	Maison				100%
12-14 ans	**École**				100%
	Maison				100%
15-16 ans	**École**				100%
	Maison				100%
17-18 ans (CÉGEP)	**École**				100%
	Maison				100%
18 ans et plus	**École**				100%
	Maison				100%

Quel âge aviez-vous lorsque vous avez commencé à communiquer à l'aide de l'anglais de façon régulière avec des locuteurs natifs de l'anglais? ________________

Quel âge aviez-vous lorsque vous avez commencé à communiquer à l'aide d'une autre langue (laquelle ? _______________) de façon régulière avec des locuteurs natifs de cette langue? _______________

Utilisiez-vous l'anglais avec des membres de votre famille? _______________

a) Si oui, avec qui (mère, père, frère, sœur, etc) ? _______________

Utilisiez-vous une autre langue avec des membres de votre famille? _______________

a) Si oui, avec qui (mère, père, frère, sœur, etc) ? ________________

Indiquez, en pourcentage, votre exposition quotidienne actuelle à chacune de ces langues dans chacune des situations suivantes :

	Français	**Anglais**	**Autre : _____**	**Total**
Amis				100%
Conjoint(e)				100%
Membres de la famille				100%
École/Travail				100%
Temps libres (sport, hobby, etc)				100%
TV/Radio				100%
Lecture (livres, journaux, Internet)				100%

Quelle langue considérez-vous comme votre langue dominante ?

Français ☐ Anglais ☐ Français et anglais de façon égale ☐ Français et autre (________) de façon égale ☐

Avez-vous suivi des cours d'anglais après 18 ans ? Oui ☐ Non ☐

Si oui : Quand ? ______________ Pour combien de temps ? ______________

Niveau ? ______________

Suivez-vous actuellement de cours d'anglais? Oui ☐ Non ☐ Niveau ? _______

Appendix 2 – LexTALE_FR: instructions and items

Test de Vocabulaire Français

Bonjour. Ceci est un test de vocabulaire français. Sur la page au verso, vous trouverez 84 séquences de lettres qui ressemblent à du français. Seulement certaines de ces séquences sont des mots français réels. Veuillez indiquer les mots que vous connaissez (ou ceux dont vous êtes convaincu(e) qu'il s'agit d'un mot français, même si vous ne pouvez pas donner leur signification exacte). Faites attention cependant, car les erreurs pénaliseront votre score. Donc, n'essayez pas d'augmenter votre score en cochant des mots que vous n'avez jamais vus.

Vocabulary Test French

Hi, this is a test of French vocabulary. On the next page you will find 84 sequences of letters that look "French". Only some of them are real words. Please, indicate the words you know (or of which you are convinced they are French words, even though you would not be able to give their precise meaning). Be careful, however: Errors are penalised. So, there is no point in trying to increase your score by adding tallies to "words" you've never seen before!

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

Appendix 3 – List of sentences

List of French correct stimuli sentences (in bold), along with the 'dummy' sentences recorded to avoid coarticulation effects after cross-splicing in the auditory modality.

#	Introductory sentences	**Experimental** (and 'dummy') sentences
1	***Marie et Jeanne jouent au hockey avec leur copain.***	***Elles osent le plaquer sur le côté.***
	Marie et Jeanne vont au marais avec leur copain.	***Elles ôtent le crapaud sur le côté.***
		Elles ôtent le placard dans la soirée.
		Elles osent le cracher dans la soirée.
2	***Les filles ne doivent pas peinturer le mur.***	***Elles osent le tapisser dans le salon.***
	Les filles doivent nettoyer le sol après la fête.	***Elles ôtent le confetti dans le salon.***
		Elles ôtent le tapis dans la soirée.
		Elles osent le confier dans la soirée.
3	***Les gardiennes ont vu Jérémy passer la barrière.***	***Elles osent le chicaner sans le maître.***
	Les assistantes ont vu le peintre terminer l'œuvre.	***Elles ôtent le chevalet sans le maître.***
		Elles ôtent le chien dans la soirée.
		Elles osent le chevaucher dans la soirée.
4	***Émilie a réalisé des films d'épouvante.***	***Elle ose les produire dans la nuit.***
	Émilie a déterré les tombeaux des aztèques.	***Elle ôte les trésors dans la nuit.***
		Elle ôte les produits dans la soirée.
		Elle ose les tremper dans la soirée.
5	***La petite écolière n'a pas peur des grands.***	***Elle ose les taquiner dans la cachette.***
	La nounou nettoie la table des jumeaux.	***Elle ôte les biberons dans la cuisine.***
		Elle ôte les tacos dans la soirée.
		Elle ose le biper dans la soirée.
6	***Les jumeaux ont volé une Cadillac rouge.***	***Ils osent la conduire dans la montagne.***
	Les frères ont ouvert la tour informatique.	***Ils ôtent la mémoire dans la machine.***
		Ils ôtent la confiture dans la soirée.
		Ils osent la mémoriser dans la soirée.
7	***Les membres de la tribu blâment la sorcière.***	***Ils osent la bannir pour la vie.***
	Les architectes du parc refont l'esquisse.	***Ils ôtent la fontaine pour la vue.***
		Ils ôtent la banane dans la soirée.
		Ils osent la fondre dans la soirée.

8	***L'artiste reçoit les vedettes de cinéma.***	***Il ose les tatouer dans le salon.***
	Michel refait les fondations de la maison.	***Il ôte les parois dans le salon.***
		Il ôte les tables dans la soirée.
		Il ose les parodier dans la soirée.
9	***Les maçonnes entament un nouveau gîte.***	***Elles sont censées le bâtir sous le pont.***
	Les armées attaquent le vieux château.	***Elles ont cassé le renfort sous le pont.***
		Elles ont cassé le bâton dans la soirée.
		Elles sont censées le renforcer dans la soirée.
10	***La secrétaire doit bien conserver les dossiers.***	***Elle est censée les stocker sur la table.***
	La fillette doit bien nettoyer les bureaux.	***Elle a cassé les flacons sur la table.***
		Elle a cassé les stops dans la soirée.
		Elle est censée les flageller dans la soirée.
11	***Julie veut rajouter les clichés à son album.***	***Elle est censée les coller dans la minute.***
	Julie doit réparer les souliers de son amie.	***Elle a cassé les talons dans la montée.***
		Elle a cassé les colliers dans la soirée.
		Elle est censée les talonner dans la soirée.
12	***Les témoins ne peuvent pas cacher la vérité.***	***Ils sont censés la révéler sous la pression.***
	Les journalistes ne filment pas l'audience.	***Ils ont cassé la caméra sous la pression.***
		Ils ont cassé la revue dans la soirée.
		Ils sont censés la camoufler dans la soirée.
13	***Les journalistes examinent la météo.***	***Ils sont censés la prédire pour le public.***
	Les délinquants nient leur méfait.	***Ils ont cassé la clôture pour le plaisir.***
		Ils ont cassé la préparation dans la soirée.
		Ils sont censés la clôturer dans la soirée.
14	***Les spectateurs ont pris une place réservée.***	***Ils sont censés la céder pour le programme.***
	Les garçons ont emprunté un jouet fragile.	***Ils ont cassé la moto pour le plaisir.***
		Ils ont cassé la scène dans la soirée.
		Ils sont censés la motiver dans la soirée.
15	***Le jardinier arrose ses tulipes.***	***Il est censé les cueillir sous la pluie.***
	Le serveur cache ses maladresses.	***Il a cassé les théières sous la pression.***
		Il a cassé les quenottes dans la soirée.
		Il est censé les théoriser dans la soirée.

16	***François aide souvent ses cousines.*** ***François lance férocement des pierres.***	***Il est censé les défendre de la milice*** ***Il a cassé les fenêtres de la maison*** *Il a cassé les défenses dans la soirée.* *Il est censé les fendre dans la soirée.*
17	***La voleuse a dérobé le tableau.*** ***La guérisseuse a trouvé l'antidote.***	***Elle souhaite le vendre dans la rue*** ***Elle soigne le peuple dans la rue*** *Elle soigne le ventre dans la soirée.* *Elle souhaite le peupler dans la soirée.*
18	***La choriste apprivoise les couplets de la chanson.*** ***L'infirmière travaille dans les services de pédiatrie.***	***Elle souhaite les chanter dans la soirée*** ***Elle soigne les bébés dans la soirée*** *Elle soigne les chanteurs dans la soirée.* *Elle souhaite les bénir dans la soirée.*
19	***L'illustratrice a inventé un nouveau personnage.*** ***L'ambulancière a trouvé le char accidenté.***	***Elle souhaite le dessiner dans le véhicule*** ***Elle soigne le passager dans le véhicule*** *Elle soigne le dessinateur dans la soirée.* *Elle souhaite le passer dans la soirée.*
20	***La demoiselle regarde les biscuits encore chauds.*** ***La chirurgienne s'occupe des patients en urgence.***	***Elle souhaite les croquer dans la minute*** ***Elle soigne les traumas dans la minute*** *Elle soigne les crocodiles dans la soirée.* *Elle souhaite les traumatiser dans la soirée.*
21	***Le marchand a reçu des pépites d'or.*** ***L'enfant joue au docteur avec ses copines.***	***Il souhaite les peser pour le fun*** ***Il soigne les poupées pour le fun*** *Il soigne les pessimistes dans la soirée.* *Il souhaite les pousser dans la soirée.*
22	***Le détective cherche une fin à l'intrigue.*** ***Le docteur aide l'accidentée de la route.***	***Il souhaite la conclure pour les proches*** ***Il soigne la victime pour les plâtres*** *Il soigne la conseillère dans la soirée.* *Il souhaite la vicier dans la soirée.*
23	***L'étudiant en dessin reproduit la Joconde.*** ***L'étudiant en médecine s'occupe du patient.***	***Il souhaite la peindre sans le modèle*** ***Il soigne la lèvre sans le médecin*** *Il soigne la pintade dans la soirée.* *Il souhaite la lever dans la soirée.*

24	***Pierre s'occupe de ses copines.*** ***Pierre s'occupe de ses écuries.***	***Il souhaite les saouler dans la nuit*** ***Il soigne les juments dans la nuit*** *Il soigne les souveraines dans la soirée.* *Il souhaite les jumeler dans la soirée.*
25	***Les bijoutières travaillent sur un petit loquet.*** ***Les écrivaines travaillent sur un nouveau scénario.***	***Elles ont failli le perdre sur la table*** ***Elles ont fini le texte sur la table*** *Elles ont fini le persil dans la soirée.* *Elles ont failli le texter dans la soirée.*
26	***Les campeuses ont pêché un gros saumon.*** ***Les filles ont créé un chef d'œuvre.***	***Elles ont failli le bouffer dans la hâte*** ***Elles ont fini le dessin dans la hâte*** *Elles ont fini le boulot dans la soirée.* *Elles ont failli le dessiner dans la soirée.*
27	***La policière a poursuivi des suspects.*** ***La cantatrice termine les spectacles.***	***Elle a failli les saisir pour de bon*** ***Elle a fini les concerts pour de bon*** *Elle a fini les serments dans la soirée.* *Elles ont failli les conserver dans la soirée.*
28	***La recrue a allumé les bougies blanches.*** ***La recrue a récité les proverbes scouts.***	***Elle a failli les souffler dans le noir*** ***Elle a fini les serments dans le noir*** *Elle a fini les sourcils dans la soirée.* *Elle a failli les sermonner dans la soirée.*
29	***Martine et Jules possèdent une Ferrari.*** ***Martine et Jules consultent une psychologue.***	***Ils ont failli la piloter pour la soirée*** ***Ils ont fini la thérapie pour la semaine*** *Ils ont fini la pilule dans la soirée.* *Ils ont failli la témoigner dans la soirée.*
30	***Les amis de Camille ont bu la sangria.*** ***Les amis de Camille ont mangé la collation.***	***Ils ont failli la verser sur la table*** ***Ils ont fini la pizza sur la table*** *Ils ont fini la verdure dans la soirée.* *Ils ont failli la piger dans la soirée.*
31	***Les cuisiniers ont une patate pourrie.*** ***Les élèves écrivent la dictée difficile.***	***Ils ont failli la frire dans la sauteuse*** ***Ils ont fini la phrase dans la souffrance*** *Ils ont fini la frittata dans la soirée.* *Ils ont failli la frapper dans la soirée.*

32	***Le cinéaste a trouvé les séquences ennuyeuses.*** ***L'artiste a trouvé des musiques entrainantes.***	***Il a failli les couper pour la sortie*** ***Il a fini les chansons pour la sortie*** *Il a fini les courses dans la soirée.* *Il a failli les chanter dans la soirée.*
33	***La poissonnière dit vérifier ses homards.*** ***La gestionnaire dit vérifier ses dossiers.***	***Elle prétend les sentir dans la matinée*** ***Elle présente les rapports dans la matinée*** *Elle présente les sens dans la soirée.* *Elle prétend les rapporter dans la soirée.*
34	***La secrétaire a reçu les impôts de la société.*** ***La stagiaire dévoile les horaires de la cour.***	***Elle prétend les trier pour le trimestre*** ***Elle présente les procès pour le trimestre*** *Elle présente les tribunaux dans la soirée.* *Elle prétend les promener dans la soirée.*
35	***La coureuse a égalé le record.*** ***L'animatrice a ouvert le festival.***	***Elle prétend le battre par la suite*** ***Elle présente le film par la suite*** *Elle présente le bateau dans la soirée.* *Elle prétend le filmer dans la soirée.*
36	***La chirurgienne demande son scalpel.*** ***La chirurgienne termine son croquis.***	***Elle prétend le vouloir pour la chirurgie*** ***Elle présente le visage pour la chirurgie*** *Elle présente le voile dans la soirée.* *Elle prétend le visualiser dans la soirée.*
37	***Le testeur découvre les liqueurs.*** ***Le recteur accueille les cohortes.***	***Il prétend les renifler pour le client*** ***Il présente les facultés pour le collège*** *Il présente les renommées dans la soirée.* *Il prétend les facturer dans la soirée.*
38	***L'élu du village proteste contre l'église.*** ***L'élu de la ville accueille la presse nationale.***	***Il prétend la démolir pour le public*** ***Il présente la capitale pour le public*** *Il présente la démonstration dans la soirée.* *Il prétend les capturer dans la soirée.*
39	***L'apprenti cuisinier doit essayer des recettes.*** ***L'agent immobilier doit contenter des clientes.***	***Il prétend les tester dans la journée*** ***Il présente les villas dans la journée*** *Il présente les tests dans la soirée.* *Il prétend les vilipender dans la soirée.*

40	***L'avocat ne comprend pas le sens de l'affaire.*** ***L'avocat rédige le plaidoyer de la prisonnière.***	***Il prétend la résoudre dans la journée*** ***Il présente la défense dans la journée*** *Il présente la réserve dans la soirée.* *Il prétend les défendre dans la soirée.*
41	***Les violonistes jouent pour le public.*** ***Les mécaniciennes travaillent pour Honda.***	***Elles vont le ravir de leur charme*** ***Elles font le moteur de leur char*** *Elles font le ravioli dans la soirée.* *Elles vont le motiver dans la soirée.*
42	***Les écolières voient un chat en danger.*** ***Les écolières crèvent les pneus du char.***	***Elles vont le secourir dans la hâte*** ***Elles font le sabotage dans la hâte*** *Elles font le secteur dans la soirée.* *Elles vont le saboter dans la soirée.*
43	***La danseuse envoûte toujours le spectateur.*** ***La graphiste contente toujours le client.***	***Elle va le séduire pour le plaisir*** ***Elle fait le tatouage pour le plaisir*** *Elle fait le sédatif dans la soirée.* *Elle va le tatouer dans la soirée.*
44	***Justine a ramassé des gros citrons.*** ***Justine a travaillé sur ses dossiers.***	***Elle va les presser pour le repas*** ***Elle fait les projets pour le rapport*** *Elle fait les préceptes dans la soirée.* *Elle va les proposer dans la soirée.*
45	***Les étudiants veulent écouter la partie de hockey.*** ***Les étudiants veulent faire partie de l'association.***	***Ils vont la capter dans le local*** ***Ils font la corvée dans le local*** *Ils font la capuche dans la soirée.* *Ils vont la corrompre dans la soirée.*
46	***Joël et Pierre achètent une dinde de Noël.*** ***Joël et Pierre travaillent l'adaptation du rôle.***	***Ils vont la cuisiner dans le four*** ***Ils font la comédie dans le film*** *Ils font la cuisine dans la soirée.* *Il va les posséder dans la soirée.*
47	***Le chanteur présente ses dernières tounes.*** ***Le mage présente ses dernières recettes.***	***Il va les ruiner pour la foule*** ***Il fait les potions pour la foule*** *Il fait les ruines dans la soirée.* *Ils vont la positionner dans la soirée.*

48	***Martin a quarante ans aujourd'hui.*** ***Martin dessine les sagas Star Wars aujourd'hui.***	***Il va les célébrer pour le fun*** ***Il fait les galaxies pour le fun*** *Il fait les célébrations dans la soirée.* *Il va les galvaniser dans la soirée.*
49	***Les pilleuses veulent protéger le trésor.*** ***Les pilleuses veulent trouver le trésor.***	***Elles croient le cacher dans le sable*** ***Elles creusent le tombeau dans le sable*** *Elles creusent le cachot dans la soirée.* *Elles croient le tonifier dans la soirée.*
50	***Lise fait des mauvais rêves avec des zombies.*** ***Lise fait des petits fossés pour évacuer l'eau.***	***Elle croit les tuer pour la foule*** ***Elle creuse les sillons pour la fuite*** *Elle creuse les tunnels dans la soirée.* *Elle croit les sillonner dans la soirée.*
51	***L'écolière attrape les bonbons Haribo.*** ***L'ouvrière aide les villages inondés.***	***Elle croit les voler pour la fille*** ***Elle creuse les canaux pour la foule*** *Elle creuse les volcans dans la soirée.* *Elle croit les canaliser dans la soirée.*
52	***La fillette marche sur les tapis.*** ***La pelleteuse commence les travaux.***	***Elle croit les salir de ses pieds*** ***Elle creuse les bunkers de sa pelle*** *Elle creuse les salons dans la soirée.* *Elle croit les bonifier dans la soirée.*
53	***Les garçons volent la bande-dessinée.*** ***Les croque-morts préparent les funérailles.***	***Ils croient la rendre dans la journée*** ***Ils creusent la fosse dans la journée*** *Ils creusent la remise dans la soirée.* *Ils croient la faucher dans la soirée.*
54	***Les voyageurs ont gouté une soupe infecte.*** ***Les ouvriers ont commencé l'installation.***	***Ils croient la vomir dans le jardin*** ***Ils creusent la piscine dans le jardin*** *Ils creusent la voie dans la soirée.* *Ils croient la piler dans la soirée.*
55	***Les soldats s'exercent avec la cible.*** ***Les soldats gardent leur position.***	***Ils croient la braquer pour le plaisir*** ***Ils creusent la tranchée pour le passage*** *Ils creusent la branche dans la soirée.* *Ils croient la transpercer dans la soirée.*

56	***Paul raconte des histoires à ses enfants.*** ***Paul accomplit plusieurs tâches au cimetière.***	***Il croit les dire pour la postérité*** ***Il creuse les tombes pour la postérité*** *Il creuse les diamants dans la soirée.* *Il croit les tondre dans la soirée.*
57	***La fermière voit le chat affamé.*** ***La benjamine va au centre équestre.***	***Elle daigne le nourrir dans la cour*** ***Elle peigne le nourrir dans la cour*** *Elles peignent le nounours dans la soirée.* *Elles daignent le chevaucher dans la soirée.*
58	***La patronne se débarrasse de son employé.*** ***La fermière a trouvé un petit animal.***	***Elle daigne le muter dans la firme*** ***Elle peigne le muter dans la ferme*** *Elle peigne les porcs dans la soirée.* *Elles daignent le chatouiller dans la soirée.*
59	***La modiste n'aime pas ses souliers mauves.*** ***La mère n'aime pas les cheveux sales.***	***Elle daigne les porter dans la maison*** ***Elle peigne les porter dans la maison*** *Elle peigne les porcs dans la soirée.* *Elle daigne les garder dans la soirée.*
60	***La fille gâtée n'aime pas ses cadeaux.*** ***La coiffeuse termine un de ses clients.***	***Elle daigne les jeter dans le secret*** ***Elle peigne les jeter dans le silence*** *Elle peigne les jeunes dans la soirée.* *Elle daigne les chevaucher dans la soirée.*
61	***Le professeur lit une poésie de Préfontaine.*** ***Le vétérinaire inspecte l'hyène du zoo.***	***Il daigne la traduire pour le public*** ***Il peigne la traduire pour le public*** *Il peigne la traductrice dans la soirée.* *Il daigne la crier dans la soirée.*
62	***Le partisan contredit la version officielle.*** ***Le collectionneur a trouvé une poupée rare.***	***Il daigne la démentir pour le public*** ***Il peigne la démentir pour le public*** *Ils peignent la démocrate dans la soirée.* *Ils daignent la figurer dans la soirée.*
63	***Le garçon n'aime pas les pastilles roses.*** ***Le berger n'aime pas les bêtes sales.***	***Il daigne les traîner dans son sac*** ***Il peigne les troupeaux dans son champ*** *Il peigne les traitres dans la soirée.* *Il daigne les trouver dans la soirée.*

64	***Le papa prend soin des chaises de jardin.*** ***Le papa prend soin des poupées de sa fille.***	***Il daigne les rentrer dans la cabane*** ***Il peigne les Barbies dans la cabane*** *Il peigne les renards dans la soirée.* *Il daigne les barboter dans la soirée.*
65	***Justine mise sur ses pools de hockey.*** ***Justine reçoit des mots de la mafia.***	***Elle déclare les parier dans le secret*** ***Elle déchire les papiers dans le secret*** *Elle déchire les parchemins dans la soirée.* *Elle déclare les papouiller dans la soirée.*
66	***L'éditrice a terminé sa commande de manuels.*** ***La mécène a reçu des faux sans valeur.***	***Elle déclare les livrer dans la nuit*** ***Elle déchire les tableaux dans la nuit*** *Elle déchire les livres dans la soirée.* *Elle déclare les tabasser dans la soirée.*
67	***La sorcière s'occupe de son patient.*** ***La sœur veut jouer un mauvais tour.***	***Elle déclare le saigner dans le salon*** ***Elle déchire le rideau dans le salon*** *Elle déchire le cessez-le-feu dans la soirée.* *Elle déclare le ridiculiser dans la soirée.*
68	***La sorcière libère la princesse de son charme.*** ***La sorcière rate le sort de mutation.***	***Elle déclare le rompre dans la caverne*** ***Elle déchire le livre dans la colère*** *Elle déchire le roman dans la soirée.* *Elle déclare le livrer dans la soirée.*
69	***Tom est tanné d'avoir mal aux jambes.*** ***Tom est tanné de se rappeler de son ex.***	***Il déclare les lever sur le lit*** ***Il déchire les photos sur le lit*** *Il déchire les lèvres dans la soirée.* *Il déclare les photographier dans la soirée.*
70	***Simon édite les sketchs de son ami comédien.*** ***Simon déteste la forme de ses essais.***	***Il déclare les lire dans le train*** ***Il déchire les pages dans le train*** *Il déchire les livres dans la soirée.* *Il déclare les paginer dans la soirée.*
71	***Le jury écoute la déposition de la victime.*** ***L'étudiant anticipe l'accusation de plagiat.***	***Il déclare la croire pour la sécurité*** ***Il déchire la preuve pour la sécurité*** *Il déchire la croix dans la soirée.* *Il déclare la prescrire dans la soirée.*

72	***Le papa raconte la blague.*** ***Le cuisinier prépare une entrée.***	***Il déclare la conter pour le plaisir*** ***Il déchire la laitue pour le plat*** *Il déchire la consigne dans la soirée.* *Il déclare la laisser dans la soirée.*
73	***Les infirmières soignent le genou foulé.*** ***Les infirmières achètent un café brûlant.***	***Elles doivent le fléchir dans la douleur*** ***Elles boivent le breuvage dans la douleur*** *Elles boivent le flacon dans la soirée.* *Elles doivent le breveter dans la soirée.*
74	***Les méchantes chipies veulent rire de Léo.*** ***Les jeunes filles veulent guérir de leur rhume.***	***Elles doivent le railler dans la cantine*** ***Elles boivent le bouillon dans la cuisine*** *Elles boivent le raisin dans la soirée.* *Elles doivent le bouillir dans la soirée.*
75	***Léa sort les biscuits brûlants du four.*** ***Léa teste les torréfacteurs de la ville.***	***Elle doit les poser sur la table*** ***Elle boit les cafés sur la table*** *Elle boit les potions dans la soirée.* *Elle doit les cafouiller dans la soirée.*
76	***La cuisinière a reçu les grains de café.*** ***La cuisinière reçoit les colis des vergers.***	***Elle doit les moudre dans la cuisine*** ***Elle boit les cidres dans la cuisine*** *Elle boit les mousses dans la soirée.* *Elle doit les situer dans la soirée.*
77	***Les chefs préparent la dinde de Noël.*** ***Les chefs se préparent une infusion.***	***Ils doivent la rôtir dans la cuisine*** ***Ils boivent la tisane dans la cuisine*** *Ils boivent la rosée dans la soirée.* *Ils doivent la tisser dans la soirée.*
78	***Les infirmiers s'occupent d'une patiente en état de choc.*** ***Les infirmiers prennent une pause au café du coin.***	***Ils doivent la consoler dans le calme*** ***Ils boivent la limonade dans le calme*** *Ils boivent la confiture dans la soirée.* *Ils doivent la limer dans la soirée.*
79	***Le joueur voit des formations adverses.*** ***L'enfant a cru voir des potions magiques.***	***Il doit les défier pour le jeu*** ***Il boit les lotions pour le jeu*** *Il boit les dégâts dans la soirée.* *Il doit les lotir dans la soirée.*

80	***Pierre découvre des guêpes agressives.***	***Il doit les fuir dans le village***
	Pierre visite les brasseries locales.	***Il boit les bières dans le village***
		Il boit les fruits dans la soirée.
		Il doit les biaiser dans la soirée.
81	***Les grandes sœurs ont rencontré leur petit frérot.***	***Elles espèrent le choyer dans la soirée***
	Les étudiantes ont présenté le théâtre Shakespearien.	***Elles épatent le doyen dans la soirée***
		Elles épatent le chômeur dans la soirée.
		Elles espèrent le douaner dans la soirée.
82	***Les soldates ont attrapé un nouvel otage.***	***Elles espèrent le retenir dans le combat***
	Les soldates ont développé un nouveau plan.	***Elles épatent le général dans le combat***
		Elles épatent le retraité dans la soirée.
		Elles espèrent le générer dans la soirée.
83	***La gardienne console les petits enfants.***	***Elle espère les calmer pour la soirée***
	La stagiaire présente les nouveaux projets.	***Elle épate les patrons pour la soirée***
		Elle épate les calmars dans la soirée.
		Elle espère les patrouiller dans la soirée.
84	***La journaliste a filmé les sabotages.***	***Elle espère les diffuser sur la chaîne***
	L'étudiante a présenté les manifestes.	***Elle épate les députés sur la chaîne***
		Elle épate les diffuseurs dans la soirée.
		Elle espère les débuter dans la soirée.
85	***Les voisins formulent une plainte.***	***Ils espèrent la déposer pour la cause***
	Les syndiqués manifestent sans trêve.	***Ils épatent la société pour la cause***
		Ils épatent la députée dans la soirée.
		Ils espèrent la socialiser dans la soirée.
86	***Les jardiniers repèrent une chienne endormie.***	***Ils espèrent la caresser sur la patte***
	Les saltimbanques font une acrobatie survoltée.	***Ils épatent la galerie sur la piste***
		Ils épatent la cardiologue dans la soirée.
		Ils espèrent la galoper dans la soirée.
87	***Le grand-père a vécu plusieurs aventures.***	***Il espère les décrire dans la classe***
	L'enfant a trouvé les dictées très faciles.	***Il épate les tutrices dans la classe***
		Il épate les départements dans la soirée.
		Il espère les tutoyer dans la soirée.

88	***L'animateur présente les activités.*** ***L'animateur divertit les prisonnières.***	***Il espère les débuter dans la joie*** ***Il épate les détenues dans la joie*** *Il épate les débutants dans la soirée.* *Il espère les détenir dans la soirée.*
89	***Les couturières ont raté le veston.*** ***Les chanteuses ont détesté le nouvel opéra.***	***Elles jurent le découdre pour le col*** ***Elles jugent le choriste pour le chœur*** *Elles jugent le décret dans la soirée.* *Elles jurent le chorégraphier dans la soirée.*
90	***Les techniciennes ont évalué l'aqueduc.*** ***Les productrices ont critiqué le vaudeville.***	***Elles jurent le purifier pour la pluie*** ***Elles jugent le comédien pour la pièce*** *Elles jugent le puritain dans la soirée.* *Elles jurent le comprendre dans la soirée.*
91	***La secrétaire s'occupe des papiers.*** ***La procureure lit les chefs d'accusation.***	***Elle jure les trouer pour le classeur*** ***Elle juge les truands pour le conseil*** *Elle juge les trouillards dans la soirée.* *Elle jure les truquer dans la soirée.*
92	***La maman s'occupe des dossards humides.*** ***La maman voit des braquages à la télé.***	***Elle jure les sécher sur le porche*** ***Elle juge les bandits sur le poste*** *Elle juge les sénateurs dans la soirée.* *Elle jure les bannir dans la soirée.*
93	***Les parents ferment la fenêtre pendant l'orage.*** ***Les tribunaux corrigent l'affaire de corruption.***	***Ils jurent la rouvrir dans la nuit*** ***Ils jugent la décision dans la cour*** *Ils jugent la rousse dans la soirée.* *Ils jurent la décider dans la soirée.*
94	***Les élèves ont manqué des questions obligatoires.*** ***Les cinéastes critiquent une scène d'action.***	***Ils jurent les refaire dans la soirée*** ***Ils jugent la doublure dans la nuance*** *Ils jugent les réfugiés dans la soirée.* *Ils jurent la doubler dans la soirée.*
95	***L'écolier s'occupe des épreuves de biologie.*** ***Julien a examiné les poursuites judiciaires.***	***Il jure les terminer dans la classe*** ***Il juge les voleuses dans la soirée*** *Il juge les terres dans la soirée.* *Il jure les voler dans la soirée.*

96	***Le chevalier voit les armes au sol.*** ***Le juré voit entrer les meurtrières.***	***Il jure les brandir dans la terreur.*** ***Il juge les crapules dans la terreur.*** *Il juge les braves dans la soirée.* *Il jure les craquer dans la soirée.*
97	***Les enseignantes ont reçu un nouveau stagiaire.*** ***Les enseignantes ont appris leur vocabulaire.***	***Elles partent le former pour le travail.*** ***Elles parlent le jargon pour le travail.*** *Elles parlent le forain dans la soirée.* *Elles partent le jarret dans la soirée.*
98	***Les victimes dénoncent le mafieux.*** ***Les cousines se rendent à Hawaï.***	***Elles partent le trahir dans la terreur.*** ***Elles parlent le créole dans la taverne.*** *Elles parlent le thaï dans la soirée.* *Elles partent le créer dans la soirée.*
99	***Les ouvrières montent un meuble.*** ***Les adolescentes s'échangent un secret.***	***Elles partent le sabler dans la cour.*** ***Elles parlent le verlan dans la cour.*** *Elles parlent le sabir dans la soirée.* *Elles partent le vérifier dans la soirée.*
100	***Les savantes ont un mélange trop concentré.*** ***Les infirmières soignent un patient étranger.***	***Elles partent le diluer dans le labo.*** ***Elles parlent le patois dans le labo.*** *Elles parlent le dialecte dans la soirée.* *Elles partent le patrouiller dans la soirée.*
101	***Les armées voient un trou dans la muraille.*** ***Les joueuses de tennis rencontrent le journaliste.***	***Elles partent le bloquer dans la soirée.*** ***Elles parlent le français dans la soirée.*** *Elles parlent le blédard dans la soirée.* *Elles partent le franchir dans la soirée.*
102	***Les chirurgiennes ont reçu le nouveau foie.*** ***Les enseignantes ont voyagé en territoire belge.***	***Elles partent le greffer pour le patient.*** ***Elles parlent le flamand pour le plaisir.*** *Elles parlent le grec dans la soirée.* *Elles partent le flamber dans la soirée.*
103	***Les étudiantes se réunissent pour leur rapport.*** ***Les étudiantes ont appris un nouvel idiome.***	***Elles partent le rédiger pour le cours.*** ***Elles parlent le polonais pour le cours.*** *Elles parlent le réunionnais dans la soirée.* *Elles partent le polir dans la soirée.*

104	***Les sœurs ont reçu leur chiot.*** ***Les sœurs ont participé à un séjour linguistique.***	***Elles partent le cajoler par la suite.*** ***Elles parlent le mandarin par la suite.*** *Elles parlent le catalan dans la soirée.* *Elles partent le mendier dans la soirée.*
105	***La maman ours a besoin de saumons.*** ***La cuisinière a terminé ses desserts.***	***Elle est supposée les pêcher pour les petits.*** ***Elle a savouré les soufflés pour les petits.*** *Elle a savouré les pêches dans la soirée.* *Elle est supposée les souffler dans la soirée.*
106	***La cuisinière prépare du pain.*** ***La cuisinière termine un dessert.***	***Elle est supposée le beurrer par la suite.*** ***Elle a savouré le nougat par la suite.*** *Elle a savouré le beurre dans la soirée.* *Elle est supposée le nourrir dans la soirée.*
107	***La campeuse poursuit le lapin.*** ***La voisine concocte un dessert.***	***Elle est supposée le chasser dans la campagne*** ***Elle a savouré le gâteau dans la cuisine.*** *Elle a savouré le champagne dans la soirée.* *Elle est supposée le gâter dans la soirée.*
108	***Les concurrents vont saboter la toile gagnante.*** ***Les étudiants admirent leur professeur.***	***Ils sont supposés la gâcher dans le silence.*** ***Ils ont savouré la leçon dans le silence.*** *Ils ont savouré la ganache dans la soirée.* *Ils sont supposés la lessiver dans la soirée.*
109	***Les athlètes s'entraînent peu pour la médaille.*** ***Les évadés sont capturés après leur fuite.***	***Ils sont supposés la mériter pour le principe.*** ***Ils ont savouré la liberté pour le principe.*** *Ils ont savouré la merguez dans la soirée.* *Ils sont supposés la libérer dans la soirée.*
110	***Les gardiens ont retrouvé la preuve.*** ***Les garçons ont gagné la partie.***	***Ils sont supposés la détruire dans le secret.*** ***Ils ont savouré la victoire dans le secret.*** *Ils ont savouré la détente dans la soirée.* *Ils sont supposés la vivre dans la soirée.*
111	***Le couturier a reçu beaucoup de vestes.*** ***Le cuisinier a goûté beaucoup de pâtisseries.***	***Il est supposé les coudre sur la table.*** ***Il a savouré les tartes sur la table.*** *Il a savouré les courgettes dans la soirée.* *Il est supposé les tartiner dans la soirée.*

112	***Tom s'occupe des pelouses de ses voisines.*** ***Tom se goinfre des sucreries de ses sœurs.***	***Il est supposé les tondre vers le diner.*** ***Il a savouré les gaufres vers le diner.*** *Il a savouré les thons dans la soirée.* *Il est supposé les gauler dans la soirée.*
113	***Les dresseuses trouvent un husky dans la nourriture.*** ***Les filles commencent un sport dans la nature.***	***Elles pensent le punir dans la cuisine.*** ***Elles poussent le traineau dans la colline.*** *Elles poussent le punk dans la soirée.* *Elles pensent le traîner dans la soirée.*
114	***Les étudiantes voient un camarade de classe.*** ***Les étudiantes aident le déménageur en panne.***	***Elles pensent le saluer dans la rue.*** ***Elles poussent le camion dans la rue.*** *Elles poussent le saltimbanque dans la soirée.* *Elles pensent le camoufler dans la soirée.*
115	***La grand-mère attend les visiteurs.*** ***La livreuse déplace les paquetages.***	***Elle pense les guetter dans la rue.*** ***Elle pousse les colis dans la rue.*** *Elle pousse le guet dans la soirée.* *Elle pense le collecter dans la soirée.*
116	***La chef a reçu ses langoustines.*** ***La chef a reçu ses cartons.***	***Elle pense les cuire dans la cuisine.*** ***Elle pousse les boîtes dans la cuisine.*** *Elle pousse les cuistots dans la soirée.* *Elle pense les botter dans la soirée.*
117	***La modéliste achète des escarpins dispendieux.*** ***La mécanicienne reçoit des autos en panne.***	***Elle pense les choisir dans la reserve.*** ***Elle pousse les voitures dans la ruelle.*** *Elle pousse les choux dans la soirée.* *Elle pense les voiler dans la soirée.*
118	***Les locataires ont fini leur lessive.*** ***Les locataires organisent leur cuisine.***	***Ils pensent la pendre dans la garde-robe.*** ***Ils poussent la table dans le garde-manger.*** *Ils poussent la penderie dans la soirée.* *Ils pensent la tabler dans la soirée.*
119	***Les stylistes doivent changer la couleur de la mèche.*** ***Les cochers doivent sortir la diligence embourbée.***	***Ils pensent la pâlir sans le patron.*** ***Ils poussent la calèche sans le patron.*** *Ils poussent lapalissade dans la soirée.* *Ils pensent la calmer dans la soirée.*

120	***L'inspecteur voit un lien entre les agressions.*** ***Le marchand fait le ménage des rayons.***	***Il pense les relier sans le patron.*** ***Il pousse les chariots sans le patron.*** *Il pousse les religieuses dans la soirée.* *Il pense les charrier dans la soirée.*
121	***Les cuisinières ont fini le potage.*** ***Les décoratrices s'occupent du salon.***	***Elles peuvent le chauffer dans la cuisine.*** ***Elles posent le tapis dans la cuisine.*** *Elles posent le chaudron dans la soirée.* *Elles peuvent le tapisser dans la soirée.*
122	***Les jumelles ont réparé leur vélo de route.*** ***Les jumelles ont ramené un souvenir de voyage.***	***Elles peuvent le pédaler sur le bitume.*** ***Elles posent le bibelot sur le buffet.*** *Elles posent le pédalo sur l'eau* *Elles peuvent le biberonner dans la soirée.*
123	***La jeune escrimeuse voit ses adversaires.*** ***La décoratrice ajoute les détails décoratifs.***	***Elle peut les vaincre pour le plaisir.*** ***Elle pose les cadres pour le plaisir.*** *Elle pose les vins dans la soirée.* *Elle peut les cadrer dans la soirée.*
124	***La jeune étudiante mange des grosses portions.*** ***La propriétaire redécore ses salles de bains.***	***Elle peut les réduire sans mon aide.*** ***Elle pose les miroirs sans mon aide.*** *Elle pose les rédactions dans la soirée.* *Elle peut les miroiter dans la soirée.*
125	***Les campeurs boivent l'eau du lac.*** ***Les jurés commencent leur évaluation.***	***Ils peuvent la filtrer pour la sécurité.*** ***Ils posent la question pour la sécurité.*** *Ils posent la fillette dans la soirée.* *Ils peuvent la questionner dans la soirée.*
126	***Les pâtissiers préparent la brioche.*** ***Les chasseurs reviennent de la chasse.***	***Ils peuvent la pétrir sur la table.*** ***Ils posent la carcasse sur la table.*** *Ils posent la tablette dans la soirée.* *Ils peuvent la caresser dans la soirée.*
127	***Le toiletteur lave mes chiennes.*** ***Le nettoyeur prend mes tuniques.***	***Il peut les coiffer sans mon aide.*** ***Il pose les tenues sans mon aide.*** *Il pose les colis dans la soirée.* *Il peut les tenir dans la soirée.*

128	***L'électricien a installé toutes ses prises.*** ***Le jardinier termine la cour arrière.***	***Il peut les brancher pour le chauffage*** ***Il pose les statues pour le chemin*** *Il pose les branches dans la soirée.* *Il peut les statuer dans la soirée.*
129	***Les organisatrices changent le nombre d'invités.*** ***Les partisanes de la révolution préparent un coup.***	***Elles semblent le tripler sans mon aide*** ***Elles ciblent le clergé sans mon aide*** *Elles ciblent le tribunal dans la soirée.* *Elles semblent le cliquer dans la soirée.*
130	***Les caissières ont un nouveau système.*** ***Les hôtesses préparent l'argument commercial.***	***Elles semblent le parfaire pour la vente*** ***Elles ciblent le public pour la vente*** *Elles ciblent le parti dans la soirée.* *Elles semblent le vendre dans la soirée.*
131	***La vieille dame dirige les partis.*** ***La vieille dame nourrit les oiseaux.***	***Elle semble les fonder pour le plaisir*** ***Elle cible les pigeons pour le plaisir*** *Elle cible les fondateurs dans la soirée.* *Elle semble les piger dans la soirée.*
132	***La mécanicienne a inspecté les capots.*** ***L'entraîneuse a présenté les exercices.***	***Elle semble les rabattre pour les moteurs*** ***Elle cible les biceps pour les muscles*** *Elle cible les rabbins dans la soirée.* *Elle semble les bichonner dans la soirée.*
133	***Les inspecteurs ont résolu l'affaire complexe.*** ***Les médicaments combattent la maladie.***	***Ils semblent la clore pour la famille*** ***Ils ciblent la grippe pour la famille*** *Ils ciblent la cloche dans la soirée.* *Ils semblent la griffer dans la soirée.*
134	***Les hommes prient devant la nouvelle sainte.*** ***Les comprimés combattent la blessure musculaire.***	***Ils semblent la bénir pour de bon*** ***Ils ciblent la douleur pour de bon*** *Ils ciblent la bénévole dans la soirée.* *Ils semblent la doubler dans la soirée.*
135	***Le paparazzi dépose ses deux valises à l'hôtel.*** ***Le paparazzi arrive sur les collines d'Hollywood.***	***Il semble les défaire pour la soirée*** ***Il cible les vedettes pour la soirée*** *Il cible les défenses dans la soirée.* *Il semble les vider dans la soirée.*

136	***Le voyageur ouvre toutes les victuailles.*** ***Le chiffon nettoie toutes les surfaces.***	***Il semble les dévorer dans le silence*** ***Il cible les saletés dans le salon*** *Il cible les déviances dans la soirée.* *Il semble les salir dans la soirée.*
137	***Les pâtissières utilisent du beurre trop dur.*** ***Les pâtissières cherchent le réfrigérateur.***	***Elles songent le ramollir dans la cuisine*** ***Elles longent le corridor dans la cuisine*** *Elles longent le ranch dans la soirée.* *Elles songent le corriger dans la soirée.*
138	***Les adolescentes font le dessin d'un beau cardinal.*** ***Les adolescentes cherchent le cellulaire tombé du pont.***	***Elles songent le rougir sur le bord*** ***Elles longent le canal sur le bord*** *Elles longent le rouleau dans la soirée.* *Elles songent le canaliser dans la soirée.*
139	***Jessica doit résoudre ses problèmes de math.*** ***Jessica ne trouve pas les bureaux administratifs.***	***Elle songe les déduire pour la réponse*** ***Elle longe les couloirs pour la recherche*** *Elle longe les détours dans la soirée.* *Elle songe les couler dans la soirée.*
140	***L'escrimeuse a acheté des sabres.*** ***La campeuse est égarée dans les bois.***	***Elle songe les manier dans le secret*** ***Elle longe les sentiers dans le silence*** *Elle longe les manoirs dans la soirée.* *Elle songe les sentir dans la soirée.*
141	***Les touristes approchent une vache sauvage.*** ***Les touristes marchent au bord de la mer.***	***Ils songent la traire pour le plaisir*** ***Ils longent la plage pour le plaisir*** *Ils longent la tresse dans la soirée.* *Ils songent la plagier dans la soirée.*
142	***Les designers ont créé une médaille d'or.*** ***Les accros du trekking sont rendus en Chine.***	***Ils songent la jaunir dans le silence*** ***Ils longent la muraille dans le silence*** *Ils longent la jonction dans la soirée.* *Ils songent la murmurer dans la soirée.*
143	***Le campeur prépare les tasses de café.*** ***Le campeur recherche la meilleure cachette.***	***Il songe les remplir dans le silence*** ***Il longe les rivières dans le silence*** *Il longe les remparts dans la soirée.* *Il songe les rivaliser dans la soirée.*

144	***Mario possède de vieilles chaises.*** ***Mario chasse les nombreuses taupes.***	***Il songe les polir dans le jardin*** ***Il longe les cabanes dans le jardin*** *Il longe les poteaux dans la soirée.* *Il songe les cabosser dans la soirée.*
145	***Les filles ont du pain au congélateur.*** ***Les filles font un achat au supermarché.***	***Elles viennent le dégeler pour le repas*** ***Elles tiennent le salami pour le repas*** *Elles tiennent le dégonflé dans la soirée.* *Elles viennent le saluer dans la soirée.*
146	***Les protestantes s'en prennent au vitrail de l'église.*** ***Les belles-sœurs amènent un présent.***	***Elles viennent le casser pour le principe*** ***Elles tiennent le cadeau pour le petit*** *Elles tiennent le casse-croûte dans la soirée.* *Elles viennent le cadrer dans la soirée.*
147	***La gardienne porte secours aux enfants qui se noient.*** ***La gardienne porte secours aux enfants perdus.***	***Elle vient les sauver par la main*** ***Elle tient les petits par la main*** *Elle tient les sauveurs dans la soirée.* *Elle vient les pitonner dans la soirée.*
148	***La policière voit des voyous suspicieux.*** ***La festivalière vient voir des groupes connus.***	***Elle vient les fouiller dans la poche*** ***Elle tient les billets dans la poche*** *Elle tient les fouines dans la soirée.* *Elle vient les biffer dans la soirée.*
149	***Les ouvriers n'aiment pas la nouvelle assurance.*** ***Les danseurs n'aiment pas faire une arabesque.***	***Ils viennent la négocier pour le principe*** ***Ils tiennent la position pour le principe*** *Ils tiennent la négativiste dans la soirée.* *Ils viennent la positionner dans la soirée.*
150	***Les athlètes s'entrainent sur la piste de 500m.*** ***Les amoureux investissent dans une bonbonnerie.***	***Ils viennent la courir pour le plaisir*** ***Ils tiennent la boutique pour le plaisir*** *Ils tiennent la coupe dans la soirée.* *Ils viennent la boutonner dans la soirée.*
151	***Louis prépare plusieurs semences.*** ***Louis a besoin d'aides visuelles.***	***Il vient les répandre sur la terre.*** ***Il tient les lunettes sur la tête.*** *Il tient les républicaines dans la soirée.* *Il vient les lubrifier dans la soirée.*

152	***Le docteur a reçu deux patientes.*** ***Le mari attend les cartes d'embarquement.***	***Il vient les guérir pour les parents.*** ***Il tient les valises pour les parents.*** *Il tient les guépards dans la soirée.* *Il vient les valider dans la soirée.*
153	***Sophie connaît le ballet et le hip-hop.*** ***Sophie note bien les garçons et les filles.***	***Elle affirme les danser dans la classe.*** ***Elle informe les parents dans la classe.*** *Elle informe les danseurs dans la soirée.* *Elle affirme les parier dans la soirée.*
154	***La critique culinaire trouve les brocolis fades.*** ***La recrue scoute voit des animaux sauvages.***	***Elle affirme les saler pour le gout.*** ***Elle informe les cadets pour le groupe.*** *Elle informe les salvatrices dans la soirée.* *Elle affirme les cadrer dans la soirée.*
155	***La frimeuse montre les albums de dessin.*** ***La frimeuse vole les chèques des apprentis.***	***Elle affirme les colorier de son talent.*** ***Elle informe les salariés de son talent.*** *Elle informe les colonels dans la soirée.* *Elle affirme les saler dans la soirée.*
156	***L'éditrice n'aime pas ces romans.*** ***La lectrice n'aime pas ces romans.***	***Elle affirme les relire pour la critique.*** ***Elle informe les vendeurs pour la critique.*** *Elle informe les religieux dans la soirée.* *Elle affirme les vendre dans la soirée.*
157	***Le vieillard adore ses petites-filles.*** ***Le vieillard voit des tornades à l'horizon.***	***Il affirme les chérir dans la tristesse.*** ***Il informe les gamines dans la tourmente.*** *Il informe les chérubins dans la soirée.* *Il affirme les gambader dans la soirée.*
158	***L'escroc pirate des cartes électroniques.*** ***Le soldat transmet le message du président.***	***Il affirme les saboter dans la journée.*** ***Il informe les colonies dans la journée.*** *Il informe les sabotiers dans la soirée.* *Il affirme les coloniser dans la soirée.*
159	***Le garçon finit les courtes évaluations.*** ***Le garçon voit ses copines s'évanouir.***	***Il affirme les passer par lui même.*** ***Il informe les mamans par lui même.*** *Il informe les passagers dans la soirée.* *Il affirme les manipuler dans la soirée.*

160	***Le stagiaire doit terminer la facture.*** ***Le stagiaire doit terminer l'enquête.***	***Il affirme la finir par lui même.*** ***Il informe la police par lui même.*** *Il informe la finlandaise dans la soirée.* *Il affirme la politiser dans la soirée.*

Appendix 4. Priming relations between main verbs and target words

Approximately half of our transitive verbs select for a specific semantic category of nouns. For example, the verb *boire* ‘to drink’ selects a liquid: because of this, it facilitates the integration of an incoming word when it is congruent. Therefore, the differential activations elicited by ungrammatical sentences with *boire* may partly reflect semantic processing constraints. The other half of our verbs did not show this restriction (e.g., *ôter* ‘to remove’ selects for multiple potential objects). Differences in the semantic priming relation between the main verb and the noun are in part reflected in the cloze probability measure that was collected and included in the mixed-effect models as a random slope. For the reader to have a better idea of the potential ERP differences that may be elicited by these semantic relations between main verbs, we (1) assigned a binary value for verbs that either do or do not select for a specific semantic category/field, and (2) calculated the average difference for each anomalous condition minus the correct one, for “high” and “low” constraint separately.

Table A.1. List of transitive verbs and the semantic categories they select for, when applicable. “High” constraint verb are in bold.

	Transitive verb		Paired control verb		Semantic category or feature
	French	English	French	English	
1	ôter	to remove	oser	to dare to	
2	**avoir cassé**	**to have broken**	**être censé**	**to be supposed to**	**[+ fragile]**
3	**soigner**	**to heal**	**souhaiter**	**to wish to**	**illness or wound**
4	avoir fini	to have finished	avoir failli	to almost have	
5	présenter	to present	prétendre	to pretend to	
6	faire	to do	aller	to go	
7	**creuser**	**to dig**	**croire**	**to believe**	**hole**
8	**peigner**	**to comb**	**daigner**	**to deign**	**hair**
9	**déchirer**	**to tear apart**	**déclarer**	**to declare**	**paper**
10	**boire**	**to drink**	**devoir**	**to have to**	**liquid**
11	**épater**	**to impress**	**espérer**	**to hope to**	**person**
12	juger	to judge	jurer	to swear to	
13	**parler**	**to speak**	**partir**	**to go to**	**language**
14	**avoir savouré**	**to have enjoyed**	**être supposé**	**to be supposed to**	**food**
15	pousser	to push	penser	to think about	
16	poser	to put	pouvoir	to be able to	
17	cibler	to aim	sembler	to seem	
18	**longer**	**to walk along**	**songer**	**to dream to**	**long spaces**
19	tenir	to hold	venir	to come to	
20	informer	to inform	affirmer	to state	

Note: the English translations do not necessarily reflect the syntactic categories that the French verbs select.

We calculated the difference between each anomalous condition (Syntax, Semantics, and Combined) minus the correct one. Separate electrodes were selected for each effect: the N400 was measured at C2-Cz-C3 electrodes and the P600 at P2-Pz-P3 ones. In order to strip most lexical-semantic aspects out of our syntactic category violations, we further split each condition into two: half of the verbs that immediately preceded the target words were more semantically constraining than the other half. While preserving our balanced design, we divided each of the three anomalous conditions between high and low semantic constraining conditions, for a total six sub-conditions. Below we present descriptive statistics of the ERP mean differences for each sub-condition.

Table A.2. Descriptive statistics for the N400 and P600 effect magnitudes in response to Syntax, Semantics, and Combined conditions, split by levels of semantic constraint (low vs. high).

		Low				High			
		Mean	SD	Min.	Max.	Mean	SD	Min.	Max.
N400	Syntax	-0.53	1.84	-4.68	2.59	-1.03	2.04	-4.98	3.20
	Semantics	-0.90	2.50	-6.19	5.06	-1.34	1.62	-4.85	2.03
	Combined	-1.11	1.93	-5.13	2.70	-1.93	2.17	-6.18	3.68
P600	Syntax	2.90	3.03	-2.33	9.98	2.72	2.18	-0.95	9.89
	Semantics	0.63	2.42	-4.73	4.68	0.56	2.21	-4.46	6.53
	Combined	3.59	2.87	-1.56	11.34	2.83	2.69	-1.86	13.64

We refrained from performing any statistical analyses on these data because they may be confounded with syntactic category, frequency, and differences in priming, as the experiment was not designed to specifically look at the issue of semantic constraint. However, we hope that the descriptive statistics will give the reader an impression of the potential effects of semantic constraint of the main verb. If constraint truly had an effect on the ERPs, we would expect to observe a *larger* N400 effect for the Syntax and Combined condition under high compared to low constraint, and a *smaller* N400 for the Semantics condition under high compared to low constraint. Numerically, “high” constraint verbs are associated with more negative values across components and conditions (i.e. the whole waveform is shifted toward negative values).

Appendix 5. Statistical outputs — Chapter 2

The following statistical outputs were submitted as Supplementary materials along with the following Manuscript: “ERP evidence against “syntax-first” approaches to sentence processing”.

Supporting information: Statistical analyses

Lauren A. Fromont

I. Behavioural data: Cumulative link mixed model output

```
## Cumulative Link Mixed Model fitted with the Laplace approximation
##
## formula: RESP ~ SYN * SEM + (Condition | Subj)
## data:    subset(scores_acc, Group == "L1")
##
##  link  threshold nobs  logLik    AIC      niter       max.grad cond.H
##  logit flexible  11481 -14839.38 29712.76 2391(11960) 1.92e-02 8.0e+02
##
## Random effects:
##  Groups Name            Variance Std.Dev. Corr
##  Subj   (Intercept)     0.6289   0.7931
##         ConditionSEM    1.0586   1.0289   -0.359
##         ConditionSYN    2.1502   1.4664   -0.325  0.863
##         ConditionSYNSEM 2.7203   1.6493   -0.302  0.903  0.995
## Number of groups:  Subj 36
##
## Coefficients:
##                   Estimate Std. Error z value Pr(>|z|)
## SYNINCOR            3.0997     0.2511  12.342   <2e-16 ***
## SEMINCOR            1.9711     0.1795  10.979   <2e-16 ***
## SYNINCOR:SEMINCOR  -1.5250     0.1550  -9.838   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Threshold coefficients:
##     Estimate Std. Error z value
## 1|2  -0.3013     0.1375  -2.191
## 2|3   1.0301     0.1382   7.453
## 3|4   2.2007     0.1395  15.775
## 4|5   3.7797     0.1419  26.636
```

II. ERP data: Mixed effects models outputs

A. ELAN: Time window 100 ms - 300 ms

The full models show no effect or interaction involving Syntax (SYN) or Semantics (SEM)

At Midline electrodes

```
## Linear mixed model fit by maximum likelihood . t-tests use
##   Satterthwaite's method [lmerModLmerTest]
## Formula: ELAN ~ SYN * SEM * Ant + (Cond | Subj) + (cCibleP_COR + Cond |
##     Item)
```

```
##    Data: FWC_Mid
##
##       AIC       BIC    logLik  deviance  df.resid
##  224572.0  224858.2 -112252.0  224504.0     33426
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -5.4947 -0.6335 -0.0058  0.6286  4.8604
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  Item     (Intercept)  2.6097  1.615
##           cCibleP_COR 10.3157  3.212     0.07
##           CondSEM      3.9575  1.989    -0.74 -0.25
##           CondSYN      4.9866  2.233    -0.68 -0.13  0.48
##           CondSYNSEM   4.2346  2.058    -0.74  0.04  0.64  0.48
##  Subj     (Intercept)  1.6123  1.270
##           CondSEM      1.4015  1.184     0.03
##           CondSYN      0.8912  0.944    -0.19  0.57
##           CondSYNSEM   1.1301  1.063    -0.17  0.46  0.36
##  Residual             46.4145  6.813
## Number of obs: 33460, groups:  Item, 160; Subj, 36
##
## Fixed effects:
##                            Estimate Std. Error         df t value
## (Intercept)               2.183e+00  2.705e-01  7.170e+01   8.068
## SYNINCOR                 -9.292e-02  2.806e-01  1.154e+02  -0.331
## SEMINCOR                 -1.421e-01  2.941e-01  9.215e+01  -0.483
## AntPost                  -6.326e-01  1.477e-01  3.256e+04  -4.283
## SYNINCOR:SEMINCOR         1.042e-01  4.150e-01  8.236e+01   0.251
## SYNINCOR:AntPost          1.096e-02  2.092e-01  3.256e+04   0.052
## SEMINCOR:AntPost          1.482e-01  2.108e-01  3.256e+04   0.703
## SYNINCOR:SEMINCOR:AntPost  8.829e-03  2.980e-01  3.256e+04   0.030
##                          Pr(>|t|)
## (Intercept)              1.17e-11 ***
## SYNINCOR                    0.741
## SEMINCOR                    0.630
## AntPost                  1.85e-05 ***
## SYNINCOR:SEMINCOR           0.802
## SYNINCOR:AntPost            0.958
## SEMINCOR:AntPost            0.482
## SYNINCOR:SEMINCOR:AntPost    0.976
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##                   (Intr) SYNINCOR SEMINCOR AntPst SYNINCOR:SEMINCOR
## SYNINCOR          -0.429
## SEMINCOR          -0.307  0.512
## AntPost           -0.273  0.263    0.251
## SYNINCOR:SEMINCOR  0.209 -0.737   -0.693   -0.178
## SYNINCOR:AP        0.193 -0.373   -0.177   -0.706  0.252
## SEMINCOR:AP        0.191 -0.184   -0.358   -0.701  0.254
## SYNINCOR:SEMINCOR: -0.135  0.262    0.254    0.496 -0.359
```

```
##                    SYNINCOR:A SEMINCOR:
## SYNINCOR
## SEMINCOR
## AntPost
## SYNINCOR:SEMINCOR
## SYNINCOR:AP
## SEMINCOR:AP         0.495
## SYNINCOR:SEMINCOR: -0.702     -0.707
```

Anova wrapper

```
## Analysis of Deviance Table (Type III Wald chisquare tests)
##
## Response: ELAN
##                Chisq Df Pr(>Chisq)
## (Intercept) 65.0972  1  7.129e-16 ***
## SYN          0.1096  1     0.7405
## SEM          0.2335  1     0.6289
## Ant         18.3472  1  1.841e-05 ***
## SYN:SEM      0.0630  1     0.8018
## SYN:Ant      0.0027  1     0.9582
## SEM:Ant      0.4940  1     0.4821
## SYN:SEM:Ant  0.0009  1     0.9764
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

R-squared (marginal and conditional)

```
## Warning: 'r.squaredGLMM' now calculates a revised statistic. See the help
## page.
```

```
##              R2m       R2c
## [1,] 0.001526027 0.1000653
```

At Lateral electrodes

```
## Linear mixed model fit by maximum likelihood . t-tests use
##   Satterthwaite's method [lmerModLmerTest]
## Formula: ELAN ~ SYN * SEM * Ant * Hemi + (Cond | Subj) + (cCibleP_COR +
##     Cond | Item)
##    Data: FWC_Lat
##
##       AIC       BIC    logLik  deviance  df.resid
##  755719.0  756125.2 -377817.5  755635.0    117068
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -6.8719 -0.6293 -0.0149  0.6097  6.1084
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  Item     (Intercept)  1.8226  1.3500
##           cCibleP_COR 25.3718  5.0370    0.29
##           CondSEM      3.0275  1.7400   -0.70 -0.17
##           CondSYN      3.4289  1.8517   -0.63 -0.15  0.47
##           CondSYNSEM   2.9497  1.7175   -0.61  0.15  0.61  0.43
```

```
##  Subj     (Intercept)  0.7822  0.8844
##           CondSEM      0.8365  0.9146   -0.03
##           CondSYN      0.5877  0.7666   -0.19  0.49
##           CondSYNSEM   0.8668  0.9310   -0.22  0.41  0.35
##  Residual             36.4957  6.0412
## Number of obs: 117110, groups:  Item, 160; Subj, 36
##
## Fixed effects:
##                                       Estimate Std. Error         df
## (Intercept)                          1.375e+00  1.926e-01  8.578e+01
## SYNINCOR                             1.257e-01  2.156e-01  1.484e+02
## SEMINCOR                            -8.385e-02  2.258e-01  1.169e+02
## AntPost                             -5.036e-01  1.000e-01  1.162e+05
## Hemiright                           -7.499e-02  9.260e-02  1.162e+05
## SYNINCOR:SEMINCOR                   -7.153e-03  3.211e-01  1.051e+02
## SYNINCOR:AntPost                    -6.548e-02  1.417e-01  1.162e+05
## SEMINCOR:AntPost                     5.775e-02  1.428e-01  1.162e+05
## SYNINCOR:Hemiright                  -2.406e-01  1.311e-01  1.162e+05
## SEMINCOR:Hemiright                   9.118e-03  1.322e-01  1.162e+05
## AntPost:Hemiright                    5.284e-01  1.415e-01  1.162e+05
## SYNINCOR:SEMINCOR:AntPost            6.380e-03  2.018e-01  1.162e+05
## SYNINCOR:SEMINCOR:Hemiright          1.640e-01  1.868e-01  1.162e+05
## SYNINCOR:AntPost:Hemiright           6.360e-02  2.003e-01  1.162e+05
## SEMINCOR:AntPost:Hemiright          -6.228e-02  2.019e-01  1.162e+05
## SYNINCOR:SEMINCOR:AntPost:Hemiright  6.904e-02  2.854e-01  1.162e+05
##                                     t value Pr(>|t|)
## (Intercept)                           7.140 2.83e-10 ***
## SYNINCOR                              0.583 0.560653
## SEMINCOR                             -0.371 0.711017
## AntPost                              -5.035 4.79e-07 ***
## Hemiright                            -0.810 0.418067
## SYNINCOR:SEMINCOR                    -0.022 0.982268
## SYNINCOR:AntPost                     -0.462 0.643889
## SEMINCOR:AntPost                      0.405 0.685795
## SYNINCOR:Hemiright                   -1.835 0.066577 .
## SEMINCOR:Hemiright                    0.069 0.944996
## AntPost:Hemiright                     3.736 0.000187 ***
## SYNINCOR:SEMINCOR:AntPost             0.032 0.974780
## SYNINCOR:SEMINCOR:Hemiright           0.878 0.380107
## SYNINCOR:AntPost:Hemiright            0.317 0.750885
## SEMINCOR:AntPost:Hemiright           -0.308 0.757709
## SYNINCOR:SEMINCOR:AntPost:Hemiright   0.242 0.808851
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Correlation matrix not shown by default, as p = 16 > 12.
## Use print(x, correlation=TRUE)  or
##   vcov(x)     if you need it
```

Anova wrapper

```
## Analysis of Deviance Table (Type III Wald chisquare tests)
##
## Response: ELAN
```

```
##                    Chisq Df Pr(>Chisq)
## (Intercept)      50.9775  1  9.343e-13 ***
## SYN               0.3401  1  0.5597675
## SEM               0.1379  1  0.7103449
## Ant              25.3510  1  4.779e-07 ***
## Hemi              0.6557  1  0.4180656
## SYN:SEM           0.0005  1  0.9822253
## SYN:Ant           0.2137  1  0.6438884
## SEM:Ant           0.1637  1  0.6857946
## SYN:Hemi          3.3655  1  0.0665744 .
## SEM:Hemi          0.0048  1  0.9449960
## Ant:Hemi         13.9553  1  0.0001872 ***
## SYN:SEM:Ant       0.0010  1  0.9747803
## SYN:SEM:Hemi      0.7704  1  0.3801048
## SYN:Ant:Hemi      0.1008  1  0.7508843
## SEM:Ant:Hemi      0.0952  1  0.7577081
## SYN:SEM:Ant:Hemi  0.0585  1  0.8088506
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

R-squared (marginal and conditional)

```
##               R2m       R2c
## [1,] 0.0008926826 0.1081204
```

B. N400: Time window 350 ms - 550 ms

The best fits involved no interactions involving the factors Syntax (SYN) and Semantics (SEM)

At Midline electrodes

```
## Linear mixed model fit by maximum likelihood . t-tests use
##   Satterthwaite's method [lmerModLmerTest]
## Formula: N400 ~ SYN + SEM + Ant + (Cond | Subj) + (cCibleP_COR + Cond |
##     Item)
##    Data: FWC_Mid
##
##       AIC       BIC    logLik  deviance  df.resid
##  236584.7  236837.2 -118262.3  236524.7     33430
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -5.1657 -0.6335 -0.0161  0.6195  4.9667
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  Item     (Intercept)  4.7060  2.1693
##           cCibleP_COR 16.3394  4.0422    0.01
##           CondSEM      7.6877  2.7727   -0.77  0.04
##           CondSYN      6.3743  2.5247   -0.75 -0.01  0.63
##           CondSYNSEM   6.2285  2.4957   -0.80 -0.03  0.66  0.74
##  Subj     (Intercept)  2.6253  1.6203
##           CondSEM      1.3324  1.1543   -0.18
```

```
##           CondSYN      0.5789  0.7609    0.19  0.10
##           CondSYNSEM   1.5737  1.2545   -0.25  0.48  0.75
##  Residual             66.5615  8.1585
## Number of obs: 33460, groups:  Item, 160; Subj, 36
##
## Fixed effects:
##              Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)     1.4645     0.3305    58.1719   4.431 4.21e-05 ***
## SYNINCOR       -0.5731     0.2227    76.3849  -2.574    0.012 *
## SEMINCOR       -1.1054     0.2124    69.1811  -5.204 1.90e-06 ***
## AntPost        -1.6588     0.0892 32502.6022 -18.596  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##          (Intr) SYNINC SEMINC
## SYNINCOR -0.228
## SEMINCOR -0.430  0.039
## AntPost  -0.135  0.000  0.000
```

Anova wrapper

```
## Analysis of Deviance Table (Type III Wald chisquare tests)
##
## Response: N400
##                Chisq Df Pr(>Chisq)
## (Intercept)  19.6333  1  9.382e-06 ***
## SYN           6.6233  1    0.01007 *
## SEM          27.0851  1  1.947e-07 ***
## Ant         345.8021  1  < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

R-squared (marginal and conditional)

```
##             R2m       R2c
## [1,] 0.01436214 0.1120388
```

At Lateral electrodes

```
## Linear mixed model fit by maximum likelihood . t-tests use
##   Satterthwaite's method [lmerModLmerTest]
## Formula: N400 ~ Hemi + SYN + SEM + Ant + (Cond | Subj) + (cCibleP_COR +
##     Cond | Item)
##    Data: FWC_Lat
##
##       AIC       BIC    logLik  deviance  df.resid
##  796058.9  796358.7 -397998.5  795996.9    117079
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -6.1593 -0.6179 -0.0094  0.6134  6.2678
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
```

```
##  Item     (Intercept)  3.3479  1.8297
##           cCibleP_COR 76.8194  8.7647    0.34
##           CondSEM      5.4151  2.3270   -0.66  0.02
##           CondSYN      4.9373  2.2220   -0.64  0.02  0.62
##           CondSYNSEM   4.2266  2.0559   -0.62  0.11  0.61  0.63
##  Subj     (Intercept)  1.6328  1.2778
##           CondSEM      0.9262  0.9624   -0.25
##           CondSYN      0.4737  0.6882    0.18  0.09
##           CondSYNSEM   1.1805  1.0865   -0.18  0.45  0.69
##  Residual            51.4817  7.1751
## Number of obs: 117110, groups:  Item, 160; Subj, 36
##
## Fixed effects:
##               Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)  1.480e+00  2.532e-01  6.015e+01   5.846 2.20e-07 ***
## Hemiright   -6.638e-01  4.193e-02  1.162e+05 -15.831  < 2e-16 ***
## SYNINCOR    -5.223e-01  1.787e-01  8.756e+01  -2.923  0.00441 **
## SEMINCOR    -8.313e-01  1.738e-01  8.374e+01  -4.784 7.28e-06 ***
## AntPost     -1.416e+00  4.237e-02  1.162e+05 -33.419  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##           (Intr) Hmrght SYNINC SEMINC
## Hemiright -0.083
## SYNINCOR  -0.142  0.000
## SEMINCOR  -0.387  0.000  0.012
## AntPost   -0.072  0.000  0.000  0.000
```

Anova wrapper

```
## Analysis of Deviance Table (Type III Wald chisquare tests)
##
## Response: N400
##                Chisq Df Pr(>Chisq)
## (Intercept)   34.180  1  5.024e-09 ***
## Hemi         250.619  1  < 2.2e-16 ***
## SYN            8.546  1   0.003463 **
## SEM           22.890  1  1.716e-06 ***
## Ant         1116.802  1  < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

R-squared (marginal and conditional)

```
##             R2m       R2c
## [1,] 0.01367494 0.1646921
```

Model with Condition and Electrode as factors (using difference waves)

```
## Linear mixed model fit by maximum likelihood . t-tests use
##   Satterthwaite's method [lmerModLmerTest]
## Formula: N400 ~ Cond * Elec + (Cond | Subj)
##    Data: N400_long
##
```

```
##      AIC      BIC   logLik deviance df.resid
##   4966.4   5211.2  -2437.2   4874.4     1466
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.7855 -0.5857  0.0203  0.5795  3.8555
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  Subj     (Intercept) 0.8472   0.9205
##           CondSYN     1.3614   1.1668   -0.72
##  Residual             1.3014   1.1408
## Number of obs: 1512, groups:  Subj, 36
##
## Fixed effects:
##                   Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)     -1.232e+00  2.443e-01  1.921e+02  -5.042 1.06e-06 ***
## CondSYN          3.803e-01  3.318e-01  2.397e+02   1.146  0.25290
## ElecC4          -1.732e-02  2.689e-01  1.440e+03  -0.064  0.94864
## ElecCz          -2.296e-01  2.689e-01  1.440e+03  -0.854  0.39335
## ElecF3           1.573e-01  2.689e-01  1.440e+03   0.585  0.55871
## ElecF4           3.330e-02  2.689e-01  1.440e+03   0.124  0.90145
## ElecF7           6.702e-01  2.689e-01  1.440e+03   2.493  0.01279 *
## ElecF8           7.688e-01  2.689e-01  1.440e+03   2.859  0.00431 **
## ElecFp1          1.519e+00  2.689e-01  1.440e+03   5.649 1.95e-08 ***
## ElecFp2          4.360e-01  2.689e-01  1.440e+03   1.621  0.10514
## ElecFpz          5.464e-01  2.689e-01  1.440e+03   2.032  0.04232 *
## ElecFz          -5.841e-02  2.689e-01  1.440e+03  -0.217  0.82805
## ElecO1           3.060e-01  2.689e-01  1.440e+03   1.138  0.25534
## ElecO2           2.535e-01  2.689e-01  1.440e+03   0.943  0.34591
## ElecOz           2.252e-01  2.689e-01  1.440e+03   0.838  0.40238
## ElecP3           8.133e-03  2.689e-01  1.440e+03   0.030  0.97587
## ElecP4          -7.343e-02  2.689e-01  1.440e+03  -0.273  0.78483
## ElecPz          -1.439e-02  2.689e-01  1.440e+03  -0.054  0.95733
## ElecT3           5.335e-01  2.689e-01  1.440e+03   1.984  0.04742 *
## ElecT4           4.881e-01  2.689e-01  1.440e+03   1.815  0.06969 .
## ElecT5           4.532e-01  2.689e-01  1.440e+03   1.686  0.09210 .
## ElecT6           1.620e-01  2.689e-01  1.440e+03   0.602  0.54698
## CondSYN:ElecC4  -1.428e-01  3.803e-01  1.440e+03  -0.376  0.70733
## CondSYN:ElecCz   2.155e-01  3.803e-01  1.440e+03   0.567  0.57093
## CondSYN:ElecF3   2.222e-01  3.803e-01  1.440e+03   0.584  0.55903
## CondSYN:ElecF4   7.218e-02  3.803e-01  1.440e+03   0.190  0.84948
## CondSYN:ElecF7   8.403e-02  3.803e-01  1.440e+03   0.221  0.82514
## CondSYN:ElecF8  -1.210e-01  3.803e-01  1.440e+03  -0.318  0.75034
## CondSYN:ElecFp1 -2.592e-01  3.803e-01  1.440e+03  -0.682  0.49550
## CondSYN:ElecFp2  6.721e-02  3.803e-01  1.440e+03   0.177  0.85974
## CondSYN:ElecFpz  2.182e-01  3.803e-01  1.440e+03   0.574  0.56620
## CondSYN:ElecFz   3.071e-01  3.803e-01  1.440e+03   0.808  0.41942
## CondSYN:ElecO1  -2.207e-01  3.803e-01  1.440e+03  -0.581  0.56166
## CondSYN:ElecO2  -1.161e-01  3.803e-01  1.440e+03  -0.305  0.76016
## CondSYN:ElecOz  -8.730e-02  3.803e-01  1.440e+03  -0.230  0.81845
## CondSYN:ElecP3  -5.935e-02  3.803e-01  1.440e+03  -0.156  0.87599
## CondSYN:ElecP4  -3.589e-02  3.803e-01  1.440e+03  -0.094  0.92482
## CondSYN:ElecPz   6.257e-02  3.803e-01  1.440e+03   0.165  0.86932
```

```
## CondSYN:ElecT3  -2.850e-01  3.803e-01  1.440e+03  -0.749  0.45375
## CondSYN:ElecT4  -2.142e-01  3.803e-01  1.440e+03  -0.563  0.57323
## CondSYN:ElecT5  -4.329e-01  3.803e-01  1.440e+03  -1.139  0.25508
## CondSYN:ElecT6  -1.964e-01  3.803e-01  1.440e+03  -0.516  0.60568
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation matrix not shown by default, as p = 42 > 12.
## Use print(x, correlation=TRUE)  or
##   vcov(x)     if you need it
```

Anova wrapper

```
## Analysis of Deviance Table (Type III Wald chisquare tests)
##
## Response: N400
##               Chisq Df Pr(>Chisq)
## (Intercept) 25.4198  1  4.612e-07 ***
## Cond         1.3135  1     0.2518
## Elec        85.0795 20  5.304e-10 ***
## Cond:Elec   10.3458 20     0.9615
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

R-squared (marginal and conditional)

```
##             R2m       R2c
## [1,] 0.07301721 0.4143312
```

Model evaluating the effect of Cloze probability at central electrodes

```
## Linear mixed model fit by maximum likelihood . t-tests use
##   Satterthwaite's method [lmerModLmerTest]
## Formula: N400 ~ SYN * SEM * cCibleP_COR + (1 | Subj) + (1 | Item)
##    Data: N400_C
##
##      AIC      BIC   logLik deviance df.resid
##  51729.4  51805.3 -25853.7  51707.4     7269
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.1686 -0.6414 -0.0022  0.6320  4.1030
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Item     (Intercept)  0.7829  0.8848
##  Subj     (Intercept)  3.5774  1.8914
##  Residual             69.6830  8.3476
## Number of obs: 7280, groups:  Item, 160; Subj, 36
##
## Fixed effects:
##                               Estimate Std. Error         df t value
## (Intercept)                    1.16890    0.37809   62.10252   3.092
## SYNINCOR                      -0.86056    0.27626 7134.26968  -3.115
## SEMINCOR                      -1.57610    0.27810 7150.67984  -5.667
```

```
## cCibleP_COR                     1.97773    0.82864 5031.40792   2.387
## SYNINCOR:SEMINCOR               0.28354    0.39336 7153.03255   0.721
## SYNINCOR:cCibleP_COR            0.08606    1.14083 7124.25459   0.075
## SEMINCOR:cCibleP_COR           -2.99549    1.14871 7133.47776  -2.608
## SYNINCOR:SEMINCOR:cCibleP_COR   0.49445    1.62440 7135.18279   0.304
##                              Pr(>|t|)
## (Intercept)                   0.00298 **
## SYNINCOR                      0.00185 **
## SEMINCOR                     1.51e-08 ***
## cCibleP_COR                   0.01704 *
## SYNINCOR:SEMINCOR             0.47105
## SYNINCOR:cCibleP_COR          0.93987
## SEMINCOR:cCibleP_COR          0.00913 **
## SYNINCOR:SEMINCOR:cCibleP_COR  0.76084
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##                   (Intr) SYNINCOR SEMINCOR cCP_CO SYNINCOR:SEMINCOR
## SYNINCOR           -0.362
## SEMINCOR           -0.360  0.494
## cCibleP_COR        -0.032  0.045    0.046
## SYNINCOR:SEMINCOR   0.255 -0.703   -0.707   -0.031
## SYNINCOR:CP         0.023 -0.076   -0.034   -0.690  0.054
## SEMINCOR:CP         0.025 -0.033   -0.075   -0.685  0.053
## SYNINCOR:SEMINCOR: -0.016  0.053    0.054    0.486 -0.078
##                   SYNINCOR:C SEMINCOR:
## SYNINCOR
## SEMINCOR
## cCibleP_COR
## SYNINCOR:SEMINCOR
## SYNINCOR:CP
## SEMINCOR:CP         0.498
## SYNINCOR:SEMINCOR: -0.703     -0.707
```

Anova wrapper

```
## Analysis of Deviance Table (Type III Wald chisquare tests)
##
## Response: N400
##                        Chisq Df Pr(>Chisq)
## (Intercept)          9.5578  1   0.001991 **
## SYN                  9.7035  1   0.001839 **
## SEM                 32.1193  1   1.45e-08 ***
## cCibleP_COR          5.6964  1   0.017000 *
## SYN:SEM              0.5196  1   0.471027
## SYN:cCibleP_COR      0.0057  1   0.939866
## SEM:cCibleP_COR      6.8001  1   0.009115 **
## SYN:SEM:cCibleP_COR  0.0927  1   0.760830
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

R-squared (marginal and conditional)

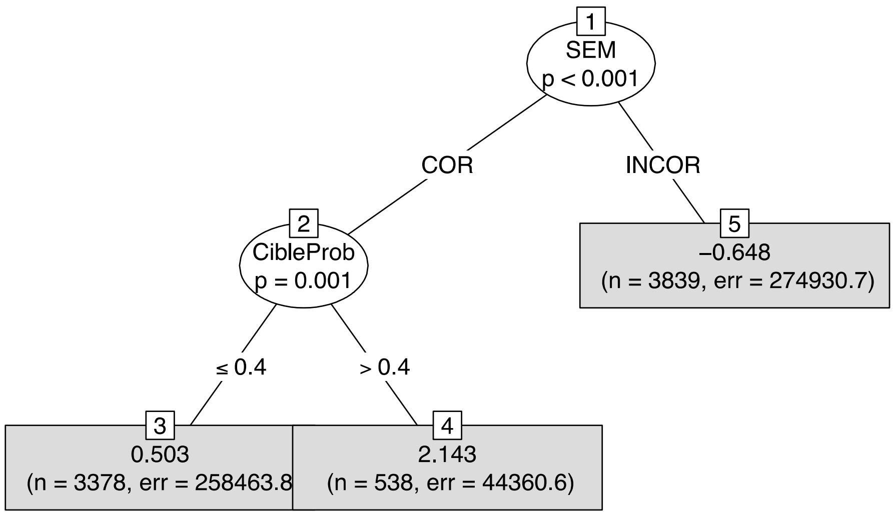
```
##             R2m        R2c
## [1,] 0.01101343 0.06925366
```

Regression tree

```
## Loading required package: grid

## Loading required package: libcoin

## Loading required package: mvtnorm
```



C. P3a: Time window 600 ms - 700 ms

At Midline electrodes

```
## Linear mixed model fit by maximum likelihood . t-tests use
##   Satterthwaite's method [lmerModLmerTest]
## Formula: eP600 ~ SYN * SEM * Ant + (Cond | Subj) + (cCibleP_COR + Cond |
##     Item)
##    Data: FWC_Mid
##
##       AIC       BIC    logLik  deviance  df.resid
##  243039.3  243325.5 -121485.7  242971.3     33426
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.4882 -0.6352 -0.0080  0.6323  4.5486
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  Item     (Intercept)  3.876   1.969
##           cCibleP_COR 18.979   4.356    -0.06
##           CondSEM      7.728   2.780    -0.73  0.04
##           CondSYN      7.707   2.776    -0.73  0.11  0.55
##           CondSYNSEM   9.699   3.114    -0.77  0.12  0.63  0.66
```

```
##  Subj     (Intercept)  2.659   1.631
##           CondSEM      1.417   1.190    -0.29
##           CondSYN      3.719   1.929     0.05  0.46
##           CondSYNSEM   5.122   2.263     0.04  0.49  0.70
##  Residual             80.576   8.976
## Number of obs: 33460, groups:  Item, 160; Subj, 36
##
## Fixed effects:
##                           Estimate Std. Error         df t value
## (Intercept)               1.812e+00  3.461e-01  6.988e+01   5.235
## SYNINCOR                  2.051e+00  4.376e-01  7.549e+01   4.687
## SEMINCOR                  1.031e-01  3.579e-01  1.208e+02   0.288
## AntPost                  -2.242e-01  1.946e-01  3.251e+04  -1.152
## SYNINCOR:SEMINCOR        -8.713e-01  5.039e-01  8.793e+01  -1.729
## SYNINCOR:AntPost         -5.071e-01  2.756e-01  3.251e+04  -1.840
## SEMINCOR:AntPost          7.684e-02  2.777e-01  3.251e+04   0.277
## SYNINCOR:SEMINCOR:AntPost  9.687e-01  3.926e-01  3.251e+04   2.467
##                          Pr(>|t|)
## (Intercept)              1.65e-06 ***
## SYNINCOR                 1.20e-05 ***
## SEMINCOR                   0.7738
## AntPost                    0.2492
## SYNINCOR:SEMINCOR          0.0873 .
## SYNINCOR:AntPost           0.0658 .
## SEMINCOR:AntPost           0.7820
## SYNINCOR:SEMINCOR:AntPost   0.0136 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##                   (Intr) SYNINCOR SEMINCOR AntPst SYNINCOR:SEMINCOR
## SYNINCOR          -0.268
## SEMINCOR          -0.489  0.486
## AntPost           -0.281  0.222    0.272
## SYNINCOR:SEMINCOR   0.317 -0.577   -0.631   -0.193
## SYNINCOR:AP         0.198 -0.315   -0.192   -0.706  0.273
## SEMINCOR:AP         0.197 -0.156   -0.388   -0.701  0.276
## SYNINCOR:SEMINCOR: -0.139  0.221    0.274    0.496 -0.390
##                   SYNINCOR:A SEMINCOR:
## SYNINCOR
## SEMINCOR
## AntPost
## SYNINCOR:SEMINCOR
## SYNINCOR:AP
## SEMINCOR:AP         0.495
## SYNINCOR:SEMINCOR: -0.702     -0.707
```

Anova wrapper

```
## Analysis of Deviance Table (Type III Wald chisquare tests)
##
## Response: eP600
##               Chisq Df Pr(>Chisq)
## (Intercept) 27.4035  1  1.651e-07 ***
## SYN         21.9710  1  2.768e-06 ***
```

```
## SEM          0.0830  1    0.77328
## Ant          1.3280  1    0.24917
## SYN:SEM      2.9903  1    0.08377 .
## SYN:Ant      3.3861  1    0.06575 .
## SEM:Ant      0.0765  1    0.78204
## SYN:SEM:Ant  6.0877  1    0.01361 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

R-squared (marginal and conditional)

```
##              R2m       R2c
## [1,] 0.007679117 0.1162467
```

Follow-up models splitting by levels of Syntax

Syntax = correct

```
## Linear mixed model fit by maximum likelihood . t-tests use
##   Satterthwaite's method [lmerModLmerTest]
## Formula: eP600 ~ SEM * Ant + (Cond | Subj) + (cCibleP_COR + Cond | Item)
##    Data: subset(FWC_Mid, SYN == "COR")
##
##      AIC      BIC   logLik deviance df.resid
## 121082.4 121190.5 -60527.2 121054.4    16706
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.5615 -0.6290 -0.0081  0.6209  4.1433
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  Item     (Intercept)  3.838   1.959
##           cCibleP_COR 29.931   5.471    -0.04
##           CondSEM      7.805   2.794    -0.75 -0.09
##  Subj     (Intercept)  2.684   1.638
##           CondSEM      1.464   1.210    -0.28
##  Residual             78.684   8.870
## Number of obs: 16720, groups:  Item, 160; Subj, 36
##
## Fixed effects:
##                   Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)       1.862e+00  3.474e-01  6.997e+01   5.360 1.01e-06 ***
## SEMINCOR          8.348e-02  3.590e-01  1.186e+02   0.233    0.817
## AntPost          -2.242e-01  1.923e-01  1.620e+04  -1.166    0.244
## SEMINCOR:AntPost  7.684e-02  2.745e-01  1.620e+04   0.280    0.780
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) SEMINCOR AntPst
## SEMINCOR    -0.484
## AntPost     -0.277  0.268
## SEMINCOR:AP  0.194 -0.382   -0.701
```

Anova wrapper

```
## Analysis of Deviance Table (Type III Wald chisquare tests)
##
## Response: eP600
##               Chisq Df Pr(>Chisq)
## (Intercept) 28.7291  1  8.324e-08 ***
## SEM          0.0541  1     0.8161
## Ant          1.3599  1     0.2436
## SEM:Ant      0.0784  1     0.7795
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Syntax = incorrect

```
## Linear mixed model fit by maximum likelihood . t-tests use
##   Satterthwaite's method [lmerModLmerTest]
## Formula: eP600 ~ SEM * Ant + (Cond | Subj) + (cCibleP_COR + Cond | Item)
##    Data: subset(FWC_Mid, SYN == "INCOR")
##
##      AIC      BIC   logLik deviance df.resid
## 121873.2 121981.3 -60922.6 121845.2    16726
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.9333 -0.6417 -0.0085  0.6415  4.5385
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  Item     (Intercept)  3.769   1.941
##           cCibleP_COR 68.927   8.302     0.18
##           CondSYNSEM   5.914   2.432    -0.53  0.09
##  Subj     (Intercept)  6.835   2.614
##           CondSYNSEM   2.830   1.682    -0.18
##  Residual             81.348   9.019
## Number of obs: 16740, groups:  Item, 160; Subj, 36
##
## Fixed effects:
##                   Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)         3.7923     0.4862    49.2265   7.800 3.76e-10 ***
## SEMINCOR           -0.7653     0.3958    73.6598  -1.934 0.056983 .
## AntPost            -0.7313     0.1961 16113.1757  -3.730 0.000192 ***
## SEMINCOR:AntPost    1.0456     0.2789 16113.1757   3.749 0.000178 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) SEMINCOR AntPst
## SEMINCOR    -0.305
## AntPost     -0.202  0.248
## SEMINCOR:AP  0.142 -0.352   -0.703
```

Anova wrapper

```
## Analysis of Deviance Table (Type III Wald chisquare tests)
```

```
##
## Response: eP600
##              Chisq Df Pr(>Chisq)
## (Intercept) 60.8358  1  6.204e-15 ***
## SEM          3.7395  1  0.0531405 .
## Ant         13.9125  1  0.0001915 ***
## SEM:Ant     14.0585  1  0.0001772 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

At Lateral electrodes

```
## Linear mixed model fit by maximum likelihood . t-tests use
##   Satterthwaite's method [lmerModLmerTest]
## Formula: eP600 ~ SYN * SEM * Ant * Hemi + (Cond | Subj) + (cCibleP_COR +
##     Cond | Item)
##    Data: FWC_Lat
##
##       AIC       BIC    logLik  deviance  df.resid
##  818598.5  819004.6 -409257.2  818514.5    117068
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -6.6979 -0.6263 -0.0181  0.6118  5.4797
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  Item     (Intercept)  3.3384  1.8271
##           cCibleP_COR 79.6321  8.9237    0.37
##           CondSEM      6.0762  2.4650   -0.61  0.07
##           CondSYN      5.8244  2.4134   -0.67  0.03  0.56
##           CondSYNSEM   7.0932  2.6633   -0.66  0.15  0.61  0.62
##  Subj     (Intercept)  1.5179  1.2320
##           CondSEM      0.9665  0.9831   -0.23
##           CondSYN      2.5881  1.6088    0.00  0.28
##           CondSYNSEM   3.1697  1.7804   -0.06  0.49  0.66
##  Residual             62.3852  7.8984
## Number of obs: 117110, groups:  Item, 160; Subj, 36
##
## Fixed effects:
##                                   Estimate Std. Error         df
## (Intercept)                      1.617e+00  2.625e-01  8.101e+01
## SYNINCOR                         1.134e+00  3.518e-01  8.295e+01
## SEMINCOR                         1.510e-02  2.834e-01  1.486e+02
## AntPost                          2.525e-01  1.308e-01  1.162e+05
## Hemiright                       -5.799e-01  1.211e-01  1.162e+05
## SYNINCOR:SEMINCOR               -7.022e-01  3.872e-01  1.253e+02
## SYNINCOR:AntPost                -6.081e-02  1.852e-01  1.162e+05
## SEMINCOR:AntPost                 1.638e-01  1.866e-01  1.162e+05
## SYNINCOR:Hemiright               5.275e-02  1.715e-01  1.162e+05
## SEMINCOR:Hemiright               6.329e-02  1.728e-01  1.162e+05
## AntPost:Hemiright                3.297e-01  1.849e-01  1.162e+05
## SYNINCOR:SEMINCOR:AntPost        6.178e-01  2.639e-01  1.162e+05
```

```
## SYNINCOR:SEMINCOR:Hemiright          4.110e-01  2.443e-01  1.162e+05
## SYNINCOR:AntPost:Hemiright           1.881e-01  2.619e-01  1.162e+05
## SEMINCOR:AntPost:Hemiright           5.644e-02  2.640e-01  1.162e+05
## SYNINCOR:SEMINCOR:AntPost:Hemiright -8.687e-02  3.732e-01  1.162e+05
##                                     t value Pr(>|t|)
## (Intercept)                           6.163 2.63e-08 ***
## SYNINCOR                              3.223  0.00182 **
## SEMINCOR                              0.053  0.95756
## AntPost                               1.931  0.05355 .
## Hemiright                            -4.790 1.67e-06 ***
## SYNINCOR:SEMINCOR                    -1.814  0.07212 .
## SYNINCOR:AntPost                     -0.328  0.74266
## SEMINCOR:AntPost                      0.877  0.38024
## SYNINCOR:Hemiright                    0.308  0.75835
## SEMINCOR:Hemiright                    0.366  0.71419
## AntPost:Hemiright                     1.783  0.07458 .
## SYNINCOR:SEMINCOR:AntPost             2.341  0.01922 *
## SYNINCOR:SEMINCOR:Hemiright           1.682  0.09252 .
## SYNINCOR:AntPost:Hemiright            0.718  0.47266
## SEMINCOR:AntPost:Hemiright            0.214  0.83068
## SYNINCOR:SEMINCOR:AntPost:Hemiright  -0.233  0.81592
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation matrix not shown by default, as p = 16 > 12.
## Use print(x, correlation=TRUE)  or
##   vcov(x)     if you need it
```

Anova wrapper

```
## Analysis of Deviance Table (Type III Wald chisquare tests)
##
## Response: eP600
##                    Chisq Df Pr(>Chisq)
## (Intercept)      37.9803  1  7.146e-10 ***
## SYN              10.3848  1   0.001271 **
## SEM               0.0028  1   0.957491
## Ant               3.7269  1   0.053543 .
## Hemi             22.9418  1  1.670e-06 ***
## SYN:SEM           3.2894  1   0.069727 .
## SYN:Ant           0.1078  1   0.742657
## SEM:Ant           0.7699  1   0.380241
## SYN:Hemi          0.0946  1   0.758352
## SEM:Hemi          0.1341  1   0.714189
## Ant:Hemi          3.1792  1   0.074581 .
## SYN:SEM:Ant       5.4815  1   0.019219 *
## SYN:SEM:Hemi      2.8301  1   0.092514 .
## SYN:Ant:Hemi      0.5157  1   0.472662
## SEM:Ant:Hemi      0.0457  1   0.830677
## SYN:SEM:Ant:Hemi  0.0542  1   0.815924
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

R-squared (marginal and conditional)

```
##              R2m      R2c
## [1,] 0.006162623 0.153648
```

Follow-up models splitting by levels of Syntax

Syntax = correct

```
## Linear mixed model fit by maximum likelihood . t-tests use
##   Satterthwaite's method [lmerModLmerTest]
## Formula: eP600 ~ SEM * Ant + (Cond | Subj) + (cCibleP_COR + Cond | Item)
##    Data: subset(FWC_Lat, SYN == "COR")
##
##       AIC       BIC    logLik  deviance  df.resid
##  501692.1  501820.5 -250832.0  501664.1     71046
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -5.7733 -0.6099 -0.0101  0.6046  6.6736
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  Item     (Intercept)  2.9522  1.7182
##           cCibleP_COR 72.3207  8.5042    0.34
##           CondSEM      5.4465  2.3338   -0.66 -0.09
##  Subj     (Intercept)  1.3366  1.1561
##           CondSEM      0.7878  0.8876   -0.20
##  Residual             67.0521  8.1885
## Number of obs: 71060, groups:  Item, 160; Subj, 36
##
## Fixed effects:
##                    Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)       1.409e+00  2.435e-01  7.564e+01   5.785 1.54e-07 ***
## SEMINCOR         -3.010e-03  2.537e-01  1.332e+02  -0.012    0.991
## AntPost          -6.006e-02  8.625e-02  7.048e+04  -0.696    0.486
## SEMINCOR:AntPost  1.070e-01  1.231e-01  7.048e+04   0.869    0.385
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) SEMINCOR AntPst
## SEMINCOR    -0.412
## AntPost     -0.188  0.180
## SEMINCOR:AP  0.131 -0.257   -0.701
```

Anova wrapper

```
## Analysis of Deviance Table (Type III Wald chisquare tests)
##
## Response: eP600
##               Chisq Df Pr(>Chisq)
## (Intercept) 33.4698  1  7.238e-09 ***
## SEM          0.0001  1     0.9905
## Ant          0.4848  1     0.4862
## SEM:Ant      0.7559  1     0.3846
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Syntax = incorrect

```
## Linear mixed model fit by maximum likelihood . t-tests use
##   Satterthwaite's method [lmerModLmerTest]
## Formula: eP600 ~ SEM * Ant + (Cond | Subj) + (cCibleP_COR + Cond | Item)
##    Data: subset(FWC_Lat, SYN == "INCOR")
##
##       AIC       BIC    logLik  deviance  df.resid
##  506350.4  506478.9 -253161.2  506322.4     71131
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -7.1442 -0.6154 -0.0062  0.6147  5.2790
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  Item     (Intercept)   3.526   1.878
##           cCibleP_COR 140.813  11.866    0.42
##           CondSYNSEM    4.499   2.121   -0.42  0.23
##  Subj     (Intercept)   3.513   1.874
##           CondSYNSEM    2.001   1.414   -0.25
##  Residual              70.892   8.420
## Number of obs: 71145, groups:  Item, 160; Subj, 36
##
## Fixed effects:
##                    Estimate Std. Error        df t value Pr(>|t|)
## (Intercept)       2.431e+00  3.505e-01  5.269e+01   6.934 5.91e-09 ***
## SEMINCOR         -4.340e-01  3.033e-01  6.929e+01  -1.431  0.15695
## AntPost          -2.168e-01  8.894e-02  7.060e+04  -2.438  0.01479 *
## SEMINCOR:AntPost  3.385e-01  1.265e-01  7.060e+04   2.676  0.00744 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) SEMINCOR AntPst
## SEMINCOR    -0.342
## AntPost     -0.134  0.155
## SEMINCOR:AP  0.094 -0.221   -0.703
```

Anova wrapper

```
## Analysis of Deviance Table (Type III Wald chisquare tests)
##
## Response: eP600
##               Chisq Df Pr(>Chisq)
## (Intercept) 48.0866  1  4.078e-12 ***
## SEM          2.0475  1   0.152460
## Ant          5.9417  1   0.014787 *
## SEM:Ant      7.1629  1   0.007443 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

D. P600: Time window 800 ms - 1200 ms

At Midline electrodes

```
## Linear mixed model fit by maximum likelihood . t-tests use
##   Satterthwaite's method [lmerModLmerTest]
## Formula: lP600 ~ SYN * SEM * Ant + (Cond | Subj) + (cCibleP_COR + Cond |
##     Item)
##    Data: FWC_Mid
##
##       AIC       BIC    logLik  deviance  df.resid
##  233857.5  234143.7 -116894.7  233789.5     33426
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -5.2582 -0.6244 -0.0030  0.6236  6.3541
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  Item     (Intercept)  3.136   1.771
##           cCibleP_COR 14.805   3.848    -0.26
##           CondSEM      4.979   2.231    -0.67  0.29
##           CondSYN      5.388   2.321    -0.69  0.21  0.44
##           CondSYNSEM   6.686   2.586    -0.75  0.25  0.51  0.50
##  Subj     (Intercept)  2.167   1.472
##           CondSEM      2.162   1.470    -0.54
##           CondSYN      4.351   2.086    -0.36  0.79
##           CondSYNSEM   4.646   2.155    -0.33  0.68  0.91
##  Residual             61.235   7.825
## Number of obs: 33460, groups:  Item, 160; Subj, 36
##
## Fixed effects:
##                            Estimate Std. Error         df t value
## (Intercept)              -1.489e-01  3.110e-01  6.751e+01  -0.479
## SYNINCOR                  1.748e+00  4.298e-01  6.113e+01   4.067
## SEMINCOR                  5.959e-01  3.481e-01  8.100e+01   1.712
## AntPost                   9.915e-02  1.696e-01  3.257e+04   0.584
## SYNINCOR:SEMINCOR        -8.220e-01  4.679e-01  7.937e+01  -1.757
## SYNINCOR:AntPost          8.845e-01  2.402e-01  3.257e+04   3.682
## SEMINCOR:AntPost         -1.363e-01  2.421e-01  3.257e+04  -0.563
## SYNINCOR:SEMINCOR:AntPost  8.452e-01  3.423e-01  3.257e+04   2.469
##                          Pr(>|t|)
## (Intercept)              0.633710
## SYNINCOR                 0.000138 ***
## SEMINCOR                 0.090723 .
## AntPost                  0.558895
## SYNINCOR:SEMINCOR        0.082802 .
## SYNINCOR:AntPost         0.000232 ***
## SEMINCOR:AntPost         0.573388
## SYNINCOR:SEMINCOR:AntPost 0.013540 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
```

```
##                    (Intr) SYNINCOR SEMINCOR AntPst SYNINCOR:SEMINCOR
## SYNINCOR           -0.477
## SEMINCOR           -0.593  0.642
## AntPost            -0.273  0.197    0.244
## SYNINCOR:SEMINCOR   0.426 -0.665   -0.763   -0.181
## SYNINCOR:AP         0.193 -0.279   -0.172   -0.706  0.257
## SEMINCOR:AP         0.191 -0.138   -0.348   -0.701  0.259
## SYNINCOR:SEMINCOR: -0.135  0.196    0.246    0.496 -0.366
##                    SYNINCOR:A SEMINCOR:
## SYNINCOR
## SEMINCOR
## AntPost
## SYNINCOR:SEMINCOR
## SYNINCOR:AP
## SEMINCOR:AP         0.495
## SYNINCOR:SEMINCOR: -0.702     -0.707
```

Anova wrapper

```
## Analysis of Deviance Table (Type III Wald chisquare tests)
##
## Response: lP600
##               Chisq Df Pr(>Chisq)
## (Intercept)  0.2291  1  0.6321635
## SYN         16.5441  1  4.753e-05 ***
## SEM          2.9309  1  0.0868966 .
## Ant          0.3416  1  0.5588909
## SYN:SEM      3.0865  1  0.0789455 .
## SYN:Ant     13.5546  1  0.0002317 ***
## SEM:Ant      0.3171  1  0.5733837
## SYN:SEM:Ant  6.0978  1  0.0135354 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

R-squared (marginal and conditional)

```
##             R2m       R2c
## [1,] 0.01827483 0.1193736
```

Follow-up models splitting by levels of Syntax

Syntax = correct

```
## Linear mixed model fit by maximum likelihood . t-tests use
##   Satterthwaite's method [lmerModLmerTest]
## Formula: lP600 ~ SEM * Ant + (Cond | Subj) + (cCibleP_COR | Item)
##    Data: subset(FWC_Mid, SYN == "COR")
##
##      AIC      BIC   logLik deviance df.resid
## 116598.1 116683.1 -58288.0 116576.1    16709
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -5.5398 -0.6129  0.0055  0.6227  4.9069
##
```

```
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  Item     (Intercept)  1.651   1.285
##           cCibleP_COR 11.743   3.427    -0.21
##  Subj     (Intercept)  2.162   1.470
##           CondSEM      2.117   1.455    -0.55
##  Residual             60.912   7.805
## Number of obs: 16720, groups:  Item, 160; Subj, 36
##
## Fixed effects:
##                    Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)      -1.475e-01  2.949e-01  5.713e+01  -0.500   0.6189
## SEMINCOR          5.739e-01  2.987e-01  4.957e+01   1.922   0.0604 .
## AntPost           9.915e-02  1.692e-01  1.638e+04   0.586   0.5579
## SEMINCOR:AntPost -1.363e-01  2.415e-01  1.638e+04  -0.565   0.5724
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) SEMINCOR AntPst
## SEMINCOR    -0.539
## AntPost     -0.287  0.283
## SEMINCOR:AP  0.201 -0.404   -0.701
```

Anova wrapper

```
## Analysis of Deviance Table (Type III Wald chisquare tests)
##
## Response: lP600
##              Chisq Df Pr(>Chisq)
## (Intercept) 0.2501  1    0.61698
## SEM         3.6923  1    0.05467 .
## Ant         0.3434  1    0.55785
## SEM:Ant     0.3187  1    0.57237
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Syntax = incorrect

```
## Linear mixed model fit by maximum likelihood . t-tests use
##   Satterthwaite's method [lmerModLmerTest]
## Formula: lP600 ~ SEM * Ant + (Cond | Subj) + (cCibleP_COR | Item)
##    Data: subset(FWC_Mid, SYN == "COR")
##
##      AIC      BIC   logLik deviance df.resid
## 116598.1 116683.1 -58288.0 116576.1    16709
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -5.5398 -0.6129  0.0055  0.6227  4.9069
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  Item     (Intercept)  1.651   1.285
##           cCibleP_COR 11.743   3.427    -0.21
```

```
##  Subj     (Intercept)  2.162   1.470
##           CondSEM      2.117   1.455    -0.55
##  Residual             60.912   7.805
## Number of obs: 16720, groups:  Item, 160; Subj, 36
##
## Fixed effects:
##                    Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)      -1.475e-01  2.949e-01  5.713e+01  -0.500   0.6189
## SEMINCOR          5.739e-01  2.987e-01  4.957e+01   1.922   0.0604 .
## AntPost           9.915e-02  1.692e-01  1.638e+04   0.586   0.5579
## SEMINCOR:AntPost -1.363e-01  2.415e-01  1.638e+04  -0.565   0.5724
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) SEMINCOR AntPst
## SEMINCOR    -0.539
## AntPost     -0.287  0.283
## SEMINCOR:AP  0.201 -0.404   -0.701
```

Anova wrapper

```
## Analysis of Deviance Table (Type III Wald chisquare tests)
##
## Response: lP600
##              Chisq Df Pr(>Chisq)
## (Intercept) 0.2501  1    0.61698
## SEM         3.6923  1    0.05467 .
## Ant         0.3434  1    0.55785
## SEM:Ant     0.3187  1    0.57237
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Follow-up models splitting by Anteriority levels

Anterior sites

```
## Linear mixed model fit by maximum likelihood . t-tests use
##   Satterthwaite's method [lmerModLmerTest]
## Formula: lP600 ~ SYN * SEM + (Cond | Subj) + (cCibleP_COR | Item)
##    Data: subset(FWC_Mid, Ant == "Ant")
##
##      AIC      BIC   logLik deviance df.resid
## 119867.5 120006.6 -59915.8 119831.5    16712
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.9337 -0.6328 -0.0009  0.6437  5.9553
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  Item     (Intercept)  0.7023  0.8381
##           cCibleP_COR  6.3509  2.5201   -0.14
##  Subj     (Intercept)  2.5347  1.5921
```

```
##          CondSEM      2.1659  1.4717   -0.49
##          CondSYN      3.7860  1.9458   -0.27  0.77
##          CondSYNSEM   4.7964  2.1901   -0.09  0.59  0.92
##  Residual            74.0152  8.6032
## Number of obs: 16730, groups:  Item, 160; Subj, 36
##
## Fixed effects:
##                   Estimate Std. Error      df t value Pr(>|t|)
## (Intercept)        -0.1283     0.3070 39.9465  -0.418   0.6781
## SYNINCOR            1.7303     0.3769 34.4377   4.591 5.65e-05 ***
## SEMINCOR            0.5447     0.3120 36.2687   1.746   0.0893 .
## SYNINCOR:SEMINCOR  -0.8111     0.4152 32.8366  -1.953   0.0593 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) SYNINCOR SEMINC
## SYNINCOR    -0.359
## SEMINCOR    -0.523  0.683
## SYNINCOR:SE  0.506 -0.605   -0.820
```

Anova wrapper

```
## Analysis of Deviance Table (Type III Wald chisquare tests)
##
## Response: lP600
##               Chisq Df Pr(>Chisq)
## (Intercept)  0.1748  1    0.67590
## SYN         21.0788  1  4.408e-06 ***
## SEM          3.0478  1    0.08085 .
## SYN:SEM      3.8158  1    0.05077 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Posterior sites

```
## Linear mixed model fit by maximum likelihood . t-tests use
##   Satterthwaite's method [lmerModLmerTest]
## Formula: lP600 ~ SYN * SEM + (Cond | Subj) + (cCibleP_COR | Item)
##    Data: subset(FWC_Mid, Ant == "Post")
##
##      AIC      BIC   logLik deviance df.resid
## 114033.5 114172.6 -56998.8 113997.5    16712
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -5.9660 -0.6294 -0.0054  0.6272  5.2095
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  Item     (Intercept)  0.677   0.8228
##           cCibleP_COR 11.764   3.4299   -0.17
##  Subj     (Intercept)  1.812   1.3461
##           CondSEM      1.704   1.3055   -0.44
##           CondSYN      4.846   2.2014   -0.26  0.79
```

```
##          CondSYNSEM   4.801   2.1911   -0.40  0.76  0.92
##  Residual             51.969   7.2089
## Number of obs: 16730, groups:  Item, 160; Subj, 36
##
## Fixed effects:
##                   Estimate Std. Error       df t value Pr(>|t|)
## (Intercept)       -0.01626    0.26254 42.48385  -0.062    0.951
## SYNINCOR           2.60683    0.40093 34.96653   6.502  1.7e-07 ***
## SEMINCOR           0.40311    0.27100 33.87243   1.488    0.146
## SYNINCOR:SEMINCOR  0.02377    0.35295 30.10222   0.067    0.947
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##            (Intr) SYNINCOR SEMINC
## SYNINCOR    -0.329
## SEMINCOR    -0.482  0.699
## SYNINCOR:SE  0.250 -0.703   -0.795
```

Anova wrapper

```
## Analysis of Deviance Table (Type III Wald chisquare tests)
##
## Response: lP600
##               Chisq Df Pr(>Chisq)
## (Intercept)  0.0038  1     0.9506
## SYN         42.2758  1  7.927e-11 ***
## SEM          2.2127  1     0.1369
## SYN:SEM      0.0045  1     0.9463
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

At Lateral electrodes

```
## Linear mixed model fit by maximum likelihood . t-tests use
##   Satterthwaite's method [lmerModLmerTest]
## Formula: lP600 ~ SYN * SEM * Ant * Hemi + (Cond | Subj) + (cCibleP_COR |
##     Item)
##    Data: FWC_Lat
##
##       AIC       BIC    logLik  deviance  df.resid
##  791196.6  791486.7 -395568.3  791136.6    117080
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -6.8560 -0.6161 -0.0029  0.6053  6.0402
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  Item     (Intercept)  0.8505  0.9222
##           cCibleP_COR 32.9886  5.7436   0.42
##  Subj     (Intercept)  1.1237  1.0601
##           CondSEM      1.3155  1.1470   -0.60
##           CondSYN      2.4688  1.5712   -0.45  0.83
```

```
##           CondSYNSEM   2.6479  1.6272   -0.38  0.74  0.84
##  Residual             49.8496  7.0604
## Number of obs: 117110, groups:  Item, 160; Subj, 36
##
## Fixed effects:
##                                     Estimate Std. Error         df
## (Intercept)                       -2.371e-01  2.058e-01  5.830e+01
## SYNINCOR                           8.629e-01  2.841e-01  4.409e+01
## SEMINCOR                           4.542e-01  2.209e-01  5.227e+01
## AntPost                            1.371e-01  1.169e-01  1.166e+05
## Hemiright                         -1.022e-01  1.082e-01  1.166e+05
## SYNINCOR:SEMINCOR                 -5.877e-01  2.991e-01  5.183e+01
## SYNINCOR:AntPost                   1.327e+00  1.656e-01  1.166e+05
## SEMINCOR:AntPost                   5.181e-02  1.668e-01  1.166e+05
## SYNINCOR:Hemiright                 3.345e-01  1.533e-01  1.166e+05
## SEMINCOR:Hemiright                -1.157e-03  1.545e-01  1.166e+05
## AntPost:Hemiright                  1.181e-01  1.653e-01  1.166e+05
## SYNINCOR:SEMINCOR:AntPost          4.947e-01  2.359e-01  1.166e+05
## SYNINCOR:SEMINCOR:Hemiright        1.396e-01  2.184e-01  1.166e+05
## SYNINCOR:AntPost:Hemiright        -1.209e-01  2.341e-01  1.166e+05
## SEMINCOR:AntPost:Hemiright        -1.184e-01  2.359e-01  1.166e+05
## SYNINCOR:SEMINCOR:AntPost:Hemiright  7.346e-02 3.336e-01  1.166e+05
##                                   t value Pr(>|t|)
## (Intercept)                        -1.152   0.2540
## SYNINCOR                            3.037   0.0040 **
## SEMINCOR                            2.056   0.0448 *
## AntPost                             1.173   0.2410
## Hemiright                          -0.944   0.3450
## SYNINCOR:SEMINCOR                  -1.965   0.0548 .
## SYNINCOR:AntPost                    8.014 1.12e-15 ***
## SEMINCOR:AntPost                    0.311   0.7562
## SYNINCOR:Hemiright                  2.182   0.0291 *
## SEMINCOR:Hemiright                 -0.007   0.9940
## AntPost:Hemiright                   0.715   0.4749
## SYNINCOR:SEMINCOR:AntPost           2.097   0.0360 *
## SYNINCOR:SEMINCOR:Hemiright         0.639   0.5226
## SYNINCOR:AntPost:Hemiright         -0.516   0.6055
## SEMINCOR:AntPost:Hemiright         -0.502   0.6158
## SYNINCOR:SEMINCOR:AntPost:Hemiright   0.220   0.8257
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation matrix not shown by default, as p = 16 > 12.
## Use print(x, correlation=TRUE)  or
##   vcov(x)     if you need it
```

Anova wrapper

```
## Analysis of Deviance Table (Type III Wald chisquare tests)
##
## Response: lP600
##                   Chisq Df Pr(>Chisq)
## (Intercept)       1.3274  1   0.249262
## SYN               9.2259  1   0.002386 **
```

```
## SEM               4.2266  1   0.039795 *
## Ant               1.3750  1   0.240951
## Hemi              0.8918  1   0.344988
## SYN:SEM           3.8599  1   0.049453 *
## SYN:Ant          64.2236  1  1.111e-15 ***
## SEM:Ant           0.0964  1   0.756156
## SYN:Hemi          4.7628  1   0.029080 *
## SEM:Hemi          0.0001  1   0.994025
## Ant:Hemi          0.5106  1   0.474870
## SYN:SEM:Ant       4.3982  1   0.035976 *
## SYN:SEM:Hemi      0.4088  1   0.522568
## SYN:Ant:Hemi      0.2667  1   0.605538
## SEM:Ant:Hemi      0.2518  1   0.615777
## SYN:SEM:Ant:Hemi  0.0485  1   0.825691
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

R-squared (marginal and conditional)

```
##            R2m        R2c
## [1,] 0.0168678 0.09448006
```

Follow-up models splitting by levels of Syntax

Syntax = correct

```
## Linear mixed model fit by maximum likelihood . t-tests use
##   Satterthwaite's method [lmerModLmerTest]
## Formula: lP600 ~ SEM * Ant * Hemi + (Cond | Subj) + (cCibleP_COR | Item)
##    Data: subset(FWC_Lat, SYN == "COR")
##
##       AIC       BIC    logLik  deviance  df.resid
##  393280.0  393414.7 -196625.0  393250.0     58505
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -6.1263 -0.6111  0.0012  0.6045  6.3014
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  Item     (Intercept)  1.246   1.116
##           cCibleP_COR 32.628   5.712    0.10
##  Subj     (Intercept)  1.149   1.072
##           CondSEM      1.359   1.166    -0.61
##  Residual             47.896   6.921
## Number of obs: 58520, groups:  Item, 160; Subj, 36
##
## Fixed effects:
##                           Estimate Std. Error         df t value
## (Intercept)             -2.676e-01  2.166e-01  6.705e+01  -1.235
## SEMINCOR                 4.689e-01  2.227e-01  5.059e+01   2.105
## AntPost                  1.371e-01  1.146e-01  5.807e+04   1.196
## Hemiright               -1.022e-01  1.061e-01  5.807e+04  -0.963
## SEMINCOR:AntPost         5.181e-02  1.635e-01  5.807e+04   0.317
```

```
## SEMINCOR:Hemiright         -1.157e-03  1.514e-01  5.807e+04  -0.008
## AntPost:Hemiright           1.181e-01  1.620e-01  5.807e+04   0.729
## SEMINCOR:AntPost:Hemiright -1.184e-01  2.313e-01  5.807e+04  -0.512
##                            Pr(>|t|)
## (Intercept)                  0.2211
## SEMINCOR                     0.0403 *
## AntPost                      0.2316
## Hemiright                    0.3353
## SEMINCOR:AntPost             0.7514
## SEMINCOR:Hemiright           0.9939
## AntPost:Hemiright            0.4660
## SEMINCOR:AntPost:Hemiright   0.6087
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) SEMINCOR AntPst Hmrght SEMINCOR:AnP SEMINCOR:H AntP:H
## SEMINCOR     -0.557
## AntPost      -0.227  0.220
## Hemiright    -0.245  0.238    0.463
## SEMINCOR:AnP  0.159 -0.315   -0.701 -0.324
## SEMINCOR:Hm   0.172 -0.340   -0.324 -0.701  0.463
## AntPst:Hmrg   0.160 -0.156   -0.707 -0.655  0.495        0.459
## SEMINCOR:AP: -0.112  0.223    0.495  0.459 -0.707       -0.655     -0.701
```

Anova wrapper

```
## Analysis of Deviance Table (Type III Wald chisquare tests)
##
## Response: lP600
##                Chisq Df Pr(>Chisq)
## (Intercept)  1.5253  1    0.21682
## SEM          4.4316  1    0.03528 *
## Ant          1.4311  1    0.23158
## Hemi         0.9282  1    0.33533
## SEM:Ant      0.1004  1    0.75139
## SEM:Hemi     0.0001  1    0.99390
## Ant:Hemi     0.5315  1    0.46600
## SEM:Ant:Hemi 0.2621  1    0.60867
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Syntax = incorrect

```
## Linear mixed model fit by maximum likelihood . t-tests use
##   Satterthwaite's method [lmerModLmerTest]
## Formula: lP600 ~ SEM * Ant + (Cond | Subj) + (cCibleP_COR | Item)
##    Data: subset(FWC_Lat, SYN == "INCOR")
##
##       AIC       BIC    logLik  deviance  df.resid
##  493601.5  493702.4 -246789.7  493579.5     71134
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -6.2198 -0.6006  0.0067  0.6089  6.6195
```

```
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  Item     (Intercept)   1.7604  1.3268
##           cCibleP_COR 106.6002 10.3247  0.60
##  Subj     (Intercept)   2.4324  1.5596
##           CondSYNSEM    0.9114  0.9547  -0.08
##  Residual              59.5948  7.7198
## Number of obs: 71145, groups:  Item, 160; Subj, 36
##
## Fixed effects:
##                    Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)       7.242e-01  2.827e-01  4.658e+01   2.562   0.0137 *
## SEMINCOR         -3.911e-02  1.811e-01  3.781e+01  -0.216   0.8302
## AntPost           5.619e-01  8.155e-02  7.074e+04   6.890 5.61e-12 ***
## SEMINCOR:AntPost  2.275e-01  1.160e-01  7.074e+04   1.961   0.0498 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) SEMINCOR AntPst
## SEMINCOR    -0.140
## AntPost     -0.153  0.238
## SEMINCOR:AP  0.107 -0.339   -0.703
```

Anova wrapper

```
## Analysis of Deviance Table (Type III Wald chisquare tests)
##
## Response: lP600
##               Chisq Df Pr(>Chisq)
## (Intercept)  6.5636  1    0.01041 *
## SEM          0.0466  1    0.82903
## Ant         47.4762  1  5.568e-12 ***
## SEM:Ant      3.8470  1    0.04984 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Follow-up models splitting by Anteriority levels

Anterior sites

```
## Linear mixed model fit by maximum likelihood . t-tests use
##   Satterthwaite's method [lmerModLmerTest]
## Formula: lP600 ~ SYN * SEM * Hemi + (Cond | Subj) + (cCibleP_COR | Item)
##    Data: subset(FWC_Lat, Ant == "Ant")
##
##       AIC       BIC    logLik  deviance  df.resid
##  458729.7  458930.1 -229342.8  458685.7     66898
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -6.4446 -0.6129 -0.0053  0.6081  5.7381
##
```

```
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  Item     (Intercept)  0.8007  0.8948
##           cCibleP_COR 19.8646  4.4570   0.19
##  Subj     (Intercept)  1.4111  1.1879
##           CondSEM      1.4035  1.1847   -0.50
##           CondSYN      2.0450  1.4300   -0.42  0.73
##           CondSYNSEM   2.7447  1.6567   -0.21  0.63  0.73
##  Residual             54.8388  7.4053
## Number of obs: 66920, groups:  Item, 160; Subj, 36
##
## Fixed effects:
##                              Estimate Std. Error         df t value
## (Intercept)                -2.691e-01  2.273e-01  5.218e+01  -1.184
## SYNINCOR                    8.602e-01  2.653e-01  4.225e+01   3.243
## SEMINCOR                    4.796e-01  2.295e-01  4.794e+01   2.089
## Hemiright                  -1.022e-01  1.135e-01  6.633e+04  -0.900
## SYNINCOR:SEMINCOR          -5.691e-01  3.221e-01  4.487e+01  -1.767
## SYNINCOR:Hemiright          3.345e-01  1.608e-01  6.633e+04   2.081
## SEMINCOR:Hemiright         -1.157e-03  1.620e-01  6.633e+04  -0.007
## SYNINCOR:SEMINCOR:Hemiright  1.396e-01  2.290e-01  6.633e+04   0.610
##                            Pr(>|t|)
## (Intercept)                 0.24180
## SYNINCOR                    0.00231 **
## SEMINCOR                    0.04201 *
## Hemiright                   0.36792
## SYNINCOR:SEMINCOR           0.08402 .
## SYNINCOR:Hemiright          0.03746 *
## SEMINCOR:Hemiright          0.99430
## SYNINCOR:SEMINCOR:Hemiright  0.54212
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##                   (Intr) SYNINCOR SEMINCOR Hmrght SYNINCOR:SEMINCOR
## SYNINCOR          -0.444
## SEMINCOR          -0.500  0.679
## Hemiright         -0.250  0.214    0.247
## SYNINCOR:SEMINCOR   0.475 -0.666   -0.716   -0.176
## SYNINCOR:Hm        0.176 -0.303   -0.175   -0.706  0.250
## SEMINCOR:Hm        0.175 -0.150   -0.353   -0.701  0.251
## SYNINCOR:SEMINCOR: -0.124  0.213    0.250    0.496 -0.356
##                   SYNINCOR:H SEMINCOR:
## SYNINCOR
## SEMINCOR
## Hemiright
## SYNINCOR:SEMINCOR
## SYNINCOR:Hm
## SEMINCOR:Hm        0.495
## SYNINCOR:SEMINCOR: -0.702     -0.707
```

Anova wrapper

```
## Analysis of Deviance Table (Type III Wald chisquare tests)
##
```

```
## Response: lP600
##               Chisq Df Pr(>Chisq)
## (Intercept)   1.4017  1   0.236439
## SYN          10.5151  1   0.001184 **
## SEM           4.3651  1   0.036682 *
## Hemi          0.8107  1   0.367922
## SYN:SEM       3.1224  1   0.077223 .
## SYN:Hemi      4.3295  1   0.037457 *
## SEM:Hemi      0.0001  1   0.994304
## SYN:SEM:Hemi  0.3716  1   0.542118
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Posterior sites

```
## Linear mixed model fit by maximum likelihood . t-tests use
##   Satterthwaite's method [lmerModLmerTest]
## Formula: lP600 ~ SYN * SEM * Hemi + (Cond | Subj) + (cCibleP_COR | Item)
##    Data: subset(FWC_Lat, Ant == "Post")
##
##       AIC       BIC    logLik  deviance  df.resid
##  330422.3  330616.4 -165189.2  330378.3     50168
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -6.1724 -0.6250 -0.0015  0.6143  5.6717
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  Item     (Intercept)  0.8362  0.9145
##           cCibleP_COR 29.4829  5.4298   0.18
##  Subj     (Intercept)  1.3942  1.1808
##           CondSEM      1.7838  1.3356   -0.60
##           CondSYN      4.1343  2.0333   -0.43  0.78
##           CondSYNSEM   4.2749  2.0676   -0.55  0.75  0.87
##  Residual             41.5884  6.4489
## Number of obs: 50190, groups:  Item, 160; Subj, 36
##
## Fixed effects:
##                               Estimate Std. Error         df t value
## (Intercept)                 -6.770e-02  2.276e-01  5.349e+01  -0.297
## SYNINCOR                     2.204e+00  3.586e-01  3.924e+01   6.145
## SEMINCOR                     4.738e-01  2.518e-01  4.295e+01   1.882
## Hemiright                    1.593e-02  1.141e-01  4.968e+04   0.140
## SYNINCOR:SEMINCOR           -1.381e-01  3.290e-01  4.259e+01  -0.420
## SYNINCOR:Hemiright           2.136e-01  1.616e-01  4.968e+04   1.321
## SEMINCOR:Hemiright          -1.196e-01  1.629e-01  4.968e+04  -0.734
## SYNINCOR:SEMINCOR:Hemiright  2.131e-01  2.303e-01  4.968e+04   0.925
##                             Pr(>|t|)
## (Intercept)                   0.7673
## SYNINCOR                    3.18e-07 ***
## SEMINCOR                      0.0667 .
## Hemiright                     0.8890
## SYNINCOR:SEMINCOR             0.6767
```

```
## SYNINCOR:Hemiright            0.1864
## SEMINCOR:Hemiright            0.4630
## SYNINCOR:SEMINCOR:Hemiright   0.3549
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##                   (Intr) SYNINCOR SEMINCOR Hmrght SYNINCOR:SEMINCOR
## SYNINCOR          -0.437
## SEMINCOR          -0.580  0.726
## Hemiright         -0.251  0.159    0.227
## SYNINCOR:SEMINCOR   0.330 -0.722   -0.775   -0.173
## SYNINCOR:Hm         0.177 -0.225   -0.160   -0.706  0.246
## SEMINCOR:Hm         0.176 -0.112   -0.323   -0.701  0.248
## SYNINCOR:SEMINCOR: -0.124  0.158    0.229    0.496 -0.350
##                   SYNINCOR:H SEMINCOR:
## SYNINCOR
## SEMINCOR
## Hemiright
## SYNINCOR:SEMINCOR
## SYNINCOR:Hm
## SEMINCOR:Hm         0.495
## SYNINCOR:SEMINCOR: -0.702     -0.707
```

Anova wrapper

```
## Analysis of Deviance Table (Type III Wald chisquare tests)
##
## Response: lP600
##                Chisq Df Pr(>Chisq)
## (Intercept)   0.0885  1    0.76614
## SYN          37.7588  1  8.005e-10 ***
## SEM           3.5410  1    0.05987 .
## Hemi          0.0195  1    0.88902
## SYN:SEM       0.1763  1    0.67455
## SYN:Hemi      1.7457  1    0.18642
## SEM:Hemi      0.5387  1    0.46299
## SYN:SEM:Hemi  0.8560  1    0.35486
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Follow-up models splitting by Hemisphere levels

Right hemisphere

```
## Linear mixed model fit by maximum likelihood . t-tests use
##   Satterthwaite's method [lmerModLmerTest]
## Formula: lP600 ~ SYN * SEM * Ant + (Cond | Subj) + (cCibleP_COR | Item)
##    Data: subset(FWC_Lat, Hemi == "right")
##
##       AIC       BIC    logLik  deviance  df.resid
##  396474.0  396671.5 -198215.0  396430.0     58533
##
## Scaled residuals:
```

```
##      Min       1Q  Median       3Q      Max
## -6.2159 -0.6142 -0.0027  0.6049  6.0182
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  Item     (Intercept)  0.6075  0.7794
##           cCibleP_COR 20.5839  4.5370   0.26
##  Subj     (Intercept)  1.4577  1.2073
##           CondSEM      1.6821  1.2970   -0.62
##           CondSYN      2.8656  1.6928   -0.50  0.82
##           CondSYNSEM   2.9346  1.7131   -0.47  0.74  0.85
##  Residual             50.3823  7.0980
## Number of obs: 58555, groups:  Item, 160; Subj, 36
##
## Fixed effects:
##                            Estimate Std. Error         df t value
## (Intercept)                 -0.3268     0.2260    47.7059  -1.446
## SYNINCOR                     1.1885     0.3036    39.3507   3.915
## SEMINCOR                     0.4404     0.2437    42.6849   1.808
## AntPost                      0.2552     0.1175 57989.3049   2.172
## SYNINCOR:SEMINCOR           -0.4421     0.3241    41.2720  -1.364
## SYNINCOR:AntPost             1.2058     0.1664 57989.3052   7.245
## SEMINCOR:AntPost            -0.0666     0.1677 57989.3049  -0.397
## SYNINCOR:SEMINCOR:AntPost    0.5681     0.2371 57989.3051   2.396
##                           Pr(>|t|)
## (Intercept)               0.154764
## SYNINCOR                  0.000349 ***
## SEMINCOR                  0.077734 .
## AntPost                   0.029889 *
## SYNINCOR:SEMINCOR         0.179919
## SYNINCOR:AntPost          4.38e-13 ***
## SEMINCOR:AntPost          0.691321
## SYNINCOR:SEMINCOR:AntPost 0.016583 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##                   (Intr) SYNINCOR SEMINCOR AntPst SYNINCOR:SEMINCOR
## SYNINCOR          -0.505
## SEMINCOR          -0.603  0.757
## AntPost           -0.223  0.166    0.207
## SYNINCOR:SEMINCOR  0.474 -0.750   -0.804   -0.155
## SYNINCOR:AP        0.157 -0.235   -0.146   -0.706  0.220
## SEMINCOR:AP        0.156 -0.116   -0.295   -0.701  0.222
## SYNINCOR:SEMINCOR: -0.110  0.165    0.209    0.496 -0.314
##                   SYNINCOR:A SEMINCOR:
## SYNINCOR
## SEMINCOR
## AntPost
## SYNINCOR:SEMINCOR
## SYNINCOR:AP
## SEMINCOR:AP        0.495
## SYNINCOR:SEMINCOR: -0.702     -0.707
```

Anova wrapper

```
## Analysis of Deviance Table (Type III Wald chisquare tests)
##
## Response: lP600
##              Chisq Df Pr(>Chisq)
## (Intercept)  2.0904  1    0.14823
## SYN         15.3286  1  9.034e-05 ***
## SEM          3.2671  1    0.07068 .
## Ant          4.7159  1    0.02989 *
## SYN:SEM      1.8609  1    0.17253
## SYN:Ant     52.4898  1  4.325e-13 ***
## SEM:Ant      0.1577  1    0.69132
## SYN:SEM:Ant  5.7403  1    0.01658 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Left hemisphere

```
## Linear mixed model fit by maximum likelihood . t-tests use
##   Satterthwaite's method [lmerModLmerTest]
## Formula: lP600 ~ SYN * SEM * Ant + (Cond | Subj) + (cCibleP_COR + Cond |
##     Item)
##    Data: subset(FWC_Lat, Hemi == "left")
##
##       AIC       BIC    logLik  deviance  df.resid
##  394157.6  394462.8 -197044.8  394089.6     58521
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -6.8150 -0.6142 -0.0077  0.6040  6.1845
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  Item     (Intercept)  2.1821  1.4772
##           cCibleP_COR 10.0537  3.1708   -0.22
##           CondSEM      3.7891  1.9466   -0.64  0.38
##           CondSYN      3.5316  1.8793   -0.68  0.13  0.48
##           CondSYNSEM   4.1140  2.0283   -0.70  0.35  0.47  0.48
##  Subj     (Intercept)  0.9908  0.9954
##           CondSEM      1.1841  1.0882   -0.59
##           CondSYN      2.1649  1.4713   -0.43  0.83
##           CondSYNSEM   2.5038  1.5823   -0.31  0.68  0.87
##  Residual             47.8671  6.9186
## Number of obs: 58555, groups:  Item, 160; Subj, 36
##
## Fixed effects:
##                          Estimate Std. Error         df t value
## (Intercept)            -2.952e-01  2.196e-01  7.649e+01  -1.344
## SYNINCOR                9.166e-01  3.071e-01  6.453e+01   2.985
## SEMINCOR                4.528e-01  2.599e-01  8.788e+01   1.742
## AntPost                 1.371e-01  1.145e-01  5.762e+04   1.197
## SYNINCOR:SEMINCOR      -6.319e-01  3.610e-01  8.873e+01  -1.750
## SYNINCOR:AntPost        1.327e+00  1.622e-01  5.762e+04   8.178
```

```
## SEMINCOR:AntPost           5.181e-02  1.635e-01  5.762e+04   0.317
## SYNINCOR:SEMINCOR:AntPost  4.947e-01  2.311e-01  5.762e+04   2.140
##                           Pr(>|t|)
## (Intercept)                 0.1829
## SYNINCOR                    0.0040 **
## SEMINCOR                    0.0850 .
## AntPost                     0.2314
## SYNINCOR:SEMINCOR           0.0835 .
## SYNINCOR:AntPost          2.94e-16 ***
## SEMINCOR:AntPost            0.7513
## SYNINCOR:SEMINCOR:AntPost   0.0323 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##                   (Intr) SYNINCOR SEMINCOR AntPst SYNINCOR:SEMINCOR
## SYNINCOR          -0.526
## SEMINCOR          -0.619  0.673
## AntPost           -0.224  0.160    0.189
## SYNINCOR:SEMINCOR   0.478 -0.673   -0.768   -0.136
## SYNINCOR:AP        0.158 -0.226   -0.133   -0.706  0.193
## SEMINCOR:AP        0.157 -0.112   -0.270   -0.701  0.194
## SYNINCOR:SEMINCOR: -0.111  0.159    0.191    0.496 -0.274
##                   SYNINCOR:A SEMINCOR:
## SYNINCOR
## SEMINCOR
## AntPost
## SYNINCOR:SEMINCOR
## SYNINCOR:AP
## SEMINCOR:AP        0.495
## SYNINCOR:SEMINCOR: -0.702     -0.707
```

Anova wrapper

```
## Analysis of Deviance Table (Type III Wald chisquare tests)
##
## Response: lP600
##               Chisq Df Pr(>Chisq)
## (Intercept)  1.8068  1   0.178894
## SYN          8.9095  1   0.002837 **
## SEM          3.0354  1   0.081469 .
## Ant          1.4320  1   0.231443
## SYN:SEM      3.0641  1   0.080039 .
## SYN:Ant     66.8835  1   2.88e-16 ***
## SEM:Ant      0.1004  1   0.751321
## SYN:SEM:Ant  4.5804  1   0.032340 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```