Visual processing of verb-second (V2) word order in second language acquisition:
ERP Evidence from French-Swedish successive bilinguals

Jeremy Yeaton

June 2019

Abstract

This paper lays out a pilot study conducted to examine the roles of Cross-Linguistic Influence (CLI) and cognitive resources on processing and parsing in a second language. The present study builds upon one conducted by Andersson et al. (2018) investigating behavioral and electrophysiological (EEG) responses to syntactic violations in a first or second language. The main goal of the previous study was to investigate whether the presence or absence of a syntactic feature (in this case V2 word order) would predict electrophysiological and/or behavioral responses to syntactic (word order) violations in the second language. The study found that while +/- V2 in the L1 did not seem to influence behavioral data, it did show some influence on the EEG results. The goal of the present study is twofold. We seek first to replicate the previous online and offline results found in a -V2 population, and then investigate the role of linguistic and cognitive variables in modulating the anticipated effect. We show that the same effect demonstrated by the -V2 participants in Andersson et al. (2018) is borne out in our French participants in the form of an anterior P600 effect. Furthermore, our data show a relationship between offline measures of cognitive functioning and online responses to linguistic stimuli. This work has implications for our understanding of cross-linguistic influence in L2 acquisition, as well as the role of domain general cognitive functions in processing language.

10601 total words

Final thesis: Research Master in Cognitive Sciences
École Normale Supérieure, Université Paris Descartes, EHESS

Directed by Frédéric Isel
UMR 7114 Modèles, Dynamiques, Corpus (CNRS, Université Paris Nanterre)
In collaboration with:
Maria Kihlstedt, Annika Andersson, Susan Sayehli & Marianne Gullberg

Contents

0 Front Material

0.1 Declaration of originality

Andersson et al. (2018) found a cross-linguistic influence effect in word order violations in L2. We seek to probe this effect further, investigating how proficiency, selective attention, and interference control may modulate this CLI effect in late bilinguals.

0.2 Declaration of contribution

- Definition of the research question: FI, JY
- Literature review: JY
- Choice of methodology: FI, JY + AA, SS & MG
- Programming ERP task: JY, SS
- Programming executive functioning tasks: JY
- Subject recruitment: JY, MK
- Data collection: JY
- Data (pre)processing: JY
- Data analysis: JY
- Interpretation of the results: JY, FI
- Writing the thesis: JY

0.3 Pre-registration document

,

Visual processing of verb-second (V2) word order in second language acquisition: ERP Evidence from French-Swedish successive bilinguals

Jeremy Yeaton
Directed by Frédéric Isel, Université Paris Nanterre
UMR 7114 Modèles, Dynamiques, Corpus

29 May 2019

To be defended in June. Thesis will be written in English.
Suggested reviewers: Amsili, Pascal; Gullberg, Marianne

Introduction

Background and rationale The present project is centered on second language acquisition in the field of cognitive neuroscience of language. It will address the general issue of how languages interact, i.e., Cross-Linguistic Influence (CLI). More precisely, we are interested in the transfer of linguistic information between a first (L1; here French) and a second language (L2; here Swedish). We will focus on the processing of syntactic information of word order. The justification of selecting this pair of languages (i.e. French-Swedish) with respect to the question of word order processing is motivated as follows. Typologically, Swedish, like most Germanic languages, uses a verb-second (V2) word order, meaning that the finite verb of a clause or sentence is placed in second position with a single major constituent preceding it, while French does not. Consequently, the aim of our project will be to examine how second language learners adapt their parsing strategies for processing a L2. The parsing strategies used by the bilinguals when confronted with a linguistic phenomenon, i.e. V2 word order, that does not exist in their L1, will be considered as a function of different linguistic and extra-linguistic factors. In order to approach this question, we plan to conduct an EEG pilot experiment using the paradigm of violation (presence of word order incongruities in the critical Swedish sentences), while subjects read syntactically correct (V2) or incorrect (V3) sentences.

This study represents a follow-up to Andersson, Sayehli, and Gullberg (2018) which examined English-Swedish and German-Swedish bilinguals.

Key research question: How does language background affect syntactic processing in L2?

General hypotheses: Based on the previous research using this paradigm, our hypothesis is that L2 learners will not demonstrate native-like processing of syntactic features that do not exist in their L1. We propose that this effect is modulated by language experience as well as cognitive functions, such that the better a speaker is in a language, the more attention is mobilized when confronted with an error.

Methods

Method is designed to replicate that of Andersson et al (2018).

Participants

Our subjects will be French-Swedish successive bilinguals living in France.

For the purposes of the master's thesis, we will test 12 participants between the ages of 18 and 35 who learned Swedish as adults. All participants must be right-handed native speakers of French. Participants who are left-handed, older than 35, or had significant exposure to Swedish before adulthood, or who report neurological deficits or current psychiatric medication will be excluded. If time and resources permit, we will collect 18 additional participants (for a total of 30), such that we will have a sample size equal to that of the original study.

Procedure and stimuli

Procedure We are using the same stimuli and procedure as Andersson et al (2018), with the addition of a test of executive functioning. An experimental session will be divided into the following parts:

1. Participant signs consent form in native language, and is randomly assigned to one of two groups.
2. Participant is fitted with an EEG cap while they fill out a questionnaire regarding their language background, handedness, and socioeconomic status (approximately 15 minutes).
3. Participant completes the Acceptability Judgment Task (AJT) during EEG recording. Participants are presented with sentences in the L2 which they must judge as acceptable or unacceptable, indicating their response with a button box (approximately one hour). In addition to scalp electrodes, mastoid electrodes will be used for reference, and electrodes will be placed above and below the left eye, as well as at the outer canthi of both eyes to detect eye movements for purposes of artifact rejection.
4. Participant completes a Swedish proficiency test (approximately 10 minutes).
5. Participant completes the Sentence Completion Task (SCT), selecting the correct order for the subject, verb, and object in sentences of various types in the L2 (30 minutes, timed).
6. Participant completes two tasks assessing executive functioning: Stroop & Navon.
7. Participant completes an English proficiency test (approximately 10 minutes).
8. Participant is thanked, appropriately debriefed, and remunerated for their time.

Total session time should be just under 2.5 hours.

Stimuli All linguistic stimuli deployed (i.e., those in the AJT and SCT) are those used in the Andersson et al (2018) study. Because these same stimuli were found to be valid for native Swedish speakers as well as other populations of L2 learners, we will not pre-test our stimuli.

1. **Acceptability Judgement Task (AJT):** According to their group assignment, participants will receive one of two sets of stimuli, i.e., the same set of sentences but counterbalanced so that sentences having the illegal V3 word order in one list have the legal V2 word order in the other, and vice versa. Each list will have 480 sentences of the types Grammatical V2 (160), Ungrammatical V3 (160), and Fillers (160). To control for wrap-up effects, critical sentences have a final phrase between 0-5 words. Sentences are presented word by word on the center of a computer screen (white font on black background) with each word on the screen for 300ms with an inter-stimulus interval of 200ms. Final words include full stops.

The last word is followed by a blank screen for 700ms, after which three question marks appear until the acceptability judgment is made. Triggers are sent to the EEG time-locked with the critical word, i.e., the subject—the point at which the word-order violation could first be detected.

2. **Sentence Completion Task (SCT):** Each sentence consists of a lead-in fragment followed by boxes with words or word combinations that must be put in order by ranking them from "1" to "3" so that the sentence is grammatical. In experimental sentences (60), the lead-in fragment consists of one of two adverbials. Half of the sentences have long prefields with additional prepositional modifiers. Experimental sentences are intermingled with fillers (180), consisting of four sentence types: topicalizations (90), questions (30), SVX sentences (30), and negated sentences (30). Sentences are pseudo-randomized such that no more than three sentences from the same condition can appear in series.
3. **Swedish proficiency test:** Word and Grammar subsection of Swedex targeting the B1 level of the Common European Framework of Reference for Languages.
4. **English proficiency test:** Oxford placement test 2

Measures

Critical measures

1. Performance on AJT (percent correct)
2. RTs on AJT
3. ERPs during AJT (targeting frontal P300, and posterior P600 components)
4. Performance on SCT

Control measures

1. Handedness
2. Language history
3. Socioeconomic status
4. Measure of executive functioning
5. Swedish proficiency
6. English proficiency

Predictions

Since French is -V2, we would predict the ERPs to resemble those of the native English speakers in the original Andersson et al study, i.e., less native-like than the German learners. Because their participants were deliberately matched for proficiency, that study did not find any effect of proficiency level on online responses. We predict that given a range of proficiency levels, we will see a) an overall decrease in amplitude of ERPs in all time windows, and b) a decrease in difference between conditions in the 500-700ms range (i.e., P600 response). We further predict that the P300, thought to sign involvement of attentional resources, will be modulated both by language proficiency and executive functioning.

Analyses

For the AJT data, response accuracy will be measured by computing d-prime (d') scores.

For the EEG data, data will be high-pass filtered above 0.5 Hz, and low-pass filtered below 40 Hz to reduce high-frequency noise. Data is then divided into 1,100 ms epochs of 1,000 ms post-stimulus onset, and a 100 ms pre-stimulus baseline. Visual artifact rejection will be conducted using the FieldTrip toolbox, using the 'summary' method, with follow-up reviews in the 'channel' and 'trial' methods for verification.

Following the visual artifact rejection, data will be subjected to Independent Component Analysis (ICA) using the 'runica' routine implemented by the EEGLAB toolbox. The topographies and time courses of the components output by the ICA will be visually inspected to determine which represent ocular and motor artifacts, and subsequently removed. The data will be visually inspected after component removal to ensure that no more artifacts remain.

Mean amplitude will be taken for each of the following post-stimulus time windows: 300-500, 500-700, 700-900, and 900-1000ms. We will then use a repeated measures ANOVA with the within-subject factors:

1. Word order (V2/V3)
2. Hemisphere (right/left)
3. Lateral position (lateral/medial)
4. Anterior/Posterior position (frontal/ fronto-temporal/ temporal/ central/ parietal/ occipital)

Language group (native vs. L1 French) will be the between-subjects factor.

To examine the relationship between proficiency or other variables and ERP, average difference amplitudes will be calculated (V2-V3) for each electrode in the selected time windows. We will then use Pearson's correlations to examine relationships between our difference amplitude measures and offline measures.

The native Swedish data was reprocessed on the same machine using the same pipeline as the L2 data to ensure that everything worked properly and was comparable between populations.

To further probe the ERP in the future, source reconstruction analyses will be conducted on the native Swedish data, as well as the collected L2 data in the time windows of interest. We also hope to conduct exploratory clustering and classification analyses.

Interpretation

If the Pearson's correlations confirm a relationship between offline executive functioning and proficiency tasks, and the amplitude of the ERP in the 300-500ms and 500-700ms windows respectively, then our hypotheses about proficiency and attention are verified. Both of these relationships should have positive correlations.

Expected contributions

Study (paradigm, stimuli, etc.) originally designed by Annika Andersson, Susaan Sayehli, and Marianne Gullberg, who collected baseline data from native Swedish speakers. Jeremy Yeaton will be responsible for recruitment, data collection, pre-processing, and analysis. Mathilde de Saint Leger will provide support with preparation of stimulus delivery. Frédéric Isel will provide critical guidance and support throughout the process, particularly during analysis and interpretation.

1 Introduction

There is considerable debate as to the role of a native language (L1) on the learning and processing of a second language (L2). The impact of the L1 on a learned L2 is known as Cross-Linguistic Influence (CLI; Jarvis & Pavlenko, 2008; Sharwood Smith & Kellerman, 1986; Odlin, 2012). The particular case we examine here is that of Verb-Second (V2) word order, a common feature of Germanic languages.

Andersson et al. (2018) conducted a study investigating this exact question. They sought to investigate whether the presence (+V2) or absence (-V2) of V2 word order in the L1 of proficiency-matched participants impacted their online and offline responses to word order violations. The violations in question come in the form of V3 word order (sentence 2), instead of the correct V2 word order (sentence 1). In these V3 cases, the verb appears in the third position in the sentence, a structure which is not permitted in main clauses in Swedish.

(1) *Idag läste hon tidningen*
Today read she paper.DEF
'Today she read the paper'

(2) * *Idag hon läste tidningen*
* Today she read paper.DEF
*'Today read she the paper'

Andersson et al. (2018) found that learners whose L1 was +V2 did not perform significantly better than learners whose L1 was -V2 in offline production and judgment tasks. They did find, however, that the +V2 learners showed much more native-like online electrophysiological responses than the -V2 learners. These electrophysiological responses came in the form of electroencephalographic (EEG) event-related potentials (ERPs). The ERP methodology seeks to measure how a particular stimulus, in this case linguistic violations, is reflected in the electrical activity of the brain measurable at the scalp by time-locking the EEG response to the onset of the stimulus in question and then averaging across trials by condition. EEG is an appropriate tool for these investigations because the responses in question unfold very quickly, sometimes within a few milliseconds. A drawback of EEG, however, is the poor spacial resolution. Thus, while we can tell *when* something is happening, with relatively high accuracy, it is very difficult to determine *where* it is happening with any sort of precision.

In the case of the ERPs in the Andersson et al. (2018) study, the +V2 learners differed from the native Swedish group only in the last 100ms of the analyzed epoch, whereas the -V2 learners differed both from the +V2 learner and native Swedish groups at all time windows of interest. More specifically, they found that all groups showed an increased early posterior negativity followed by a larger posterior positivity in response to word order violations relative to a baseline condition. This P600 component is, however, localized not only at posterior sites. The differences between conditions during this time window were in fact most pronounced over anterior sites. This difference between conditions demonstrates that all groups were sensitive to word order violations in Swedish. Andersson et al. (2018) proposed that the frontal positivity elicited in English learners represents a different type of processing of a syntactic structure given that it is absent from their L1, as compared to the Germans where the structure is present. They further proposed that the anterior positivity could indicate increased mobilization of attentional resources in the L2 population relative to native speakers in response to a word that is unexpected in a constrained context. It is this suggestion which motivates the investigation into cognitive resources used in the present study.

Our goal here is twofold. We seek to reproduce the effect seen in -V2 learners with a different population, and to further examine what impact cognitive functions might have on the processing of word order in an L2. We are particularly interested in selective attention and interference control, as these skills are required to select the information necessary for language use in a given situation and to reduce undue influence from other stimuli or internal representations.

2 Background

2.1 V2 word order

In language typology, word order is defined as how the subject, verb, and object (S, V, O) are organized relative to one another in the main clause of declarative sentences in a language. This ordering varies across languages and language families (Greenberg, 1963). Most of the languages in the Germanic family make use of so-called V2 word order wherein the finite verb is always the second major constituent of a main clause (Dryer, 2005). This can be realized as an SVO word order (subject, verb, object), as well as an XVS or subject-verb inversion word order, where some constituent X–such as an adverbial or prepositional phrase–holds the first position in the sentence, thus relegating the subject to the third position.

While Swedish consistently displays V2 word order, this structure is secondary to the overall dominant SVO word order (Jörgensen, 1976; Westman, 1974). In contrast to both Swedish and French, German does not display SVO as a dominant word order since this ordering is only available in main clauses without auxiliaries. V2 word order is highly prevalent in German, occurring slightly more than in Swedish (Engel, 1974; Bohnacker, 2006). Contemporary English displays only SVO word order as V2 word order is no longer productive despite existing historically. Andersson et al. (2018) exploited the historical relatedness of the three Germanic languages–Swedish, German, and English–as a case study in L2 processing and acquisition, looking specifically for cross-linguistic influence of these word ordering rules.

Contemporary French falls into the same category as English, displaying only SVO word order in finite main clauses, but is typologically not as closely related to Swedish and German as English is.

2.2 Word order processing

The tracking of word order is critical to language use and comprehension. Speakers and listeners must track the incoming signal, parse it into its component words, and then arrive at an interpretation. If the syntactic rules governing the ordering of words in a sentence are broken, however, there are a variety of effects that can be observed. The processing of so-called word order violations has been a question in the literature for several decades, since Kutas & A. Hillyard (1980) was published employing the paradigm of violation in an ERP study. Many studies since then have employed the same paradigm to investigate syntactic violations using ERP-based methodologies (Neville et al., 1991; Osterhout & Nicol, 1999; A. Friederici et al., 1996; A. D. Friederici, 2002; Hagoort, 2003; Isel et al., 2007). The online responses observed, however, do vary according to the native language of the user, supporting a hypothesis involving CLI (Zawiszewski et al., 2011). Several different ERP components have been observed in response to syntactic or morpho-syntactic violations: the Early Left Anterior Negativity (ELAN) (A. Friederici et al., 1996, 1993; Hahne & Friederici, 1999, 2002; A. Friederici et al., 2004; Steinhauer & Drury, 2012), the P600 (Osterhout & Holcomb, 1992; Osterhout & Mobley, 1995; A. Friederici et al., 1996; Hahne & Friederici, 1999; Weyerts et al., 2002), and even a biphasic N400-P600 ERP (A. Friederici & Meyer, 2004). Both the LAN and P600 components have been observed in monolinguals and early bilinguals, in addition to late bilinguals, the population in question for this study (Steinhauer & Drury, 2012; Isel et al., 2007). The P600, in particular is thought to reflect a reanalysis of syntactic information (Osterhout & Mobley, 1995), or a repair of a structural violation (Isel, 2005, 2017). In a grammaticality judgement task, Moreno et al. (2010) found a P600 component in bilinguals in response to syntactic errors in the L2. In fact, the component demonstrated greater positivity in bilinguals than monolinguals, which seems to point to bilinguals mobilizing more cognitive/ executive resources in response to syntactic errors than their monolingual counterparts. This is consistent with the findings of Andersson et al. (2018). The main components identified in the Andersson et al. (2018) study were the P600 and LAN, with a P300 peak appearing in all groups but not differentiating them. This frontal P300 component is thought to indicate involvement of attentional resources or syntactic complexity.

2.3 Bilingualism & Cognitive Functioning

Many studies over the years have identified areas where bilinguals outperform their monolingual counterparts on cognitive tasks (see Bialystok et al., 2009, for a review). Some examples of such a bilingual advantage have been identified as having smaller interference and larger facilitation effects on certain tasks than their monolingual counterparts, characterized as better executive control of perception/action processing (Bialystok et al., 2008), better metalinguistic and linguistic conflict resolution skills (Wolleb et al., 2017; Ye & Zhou, 2009b; van Herten et al., 2006), better overall executive functioning and cognitive control (Ye & Zhou, 2009a; Nee et al., 2007), and enhanced language learning skills in childhood (Hopp et al., 2019).

The so-called bilingual advantage is thought to be the result of the need for bilinguals to actively monitor and control which language they are using at any given time. There are two main theories for how this language selection takes place: either the unused language(s) is/are actively inhibited, or the used language is actively selected, and the process of switching between them is a cognitively taxing task (Swainson et al., 2001). Colzato et al. (2008) proposes a distinction between these where the former would be active inhibition or the general global suppression of the nonrelevant language, whereas the latter would be reactive inhibition or the lack of suppression of specific interfering stimuli. The proposal here being that the bilingual advantage is not due to constantly actively inhibiting the the non-used language but rather to prolonged practice at maintaining the relevant attention set. They grant, however, that such selection may still involve strong inhibition of competing items.

The neural mechanisms required to control switching between languages has been documented both in populations who use more than one spoken language, as well as in speech-sign bilinguals, however it does not appear as though speech-sign bilinguals enjoy the same bilingual advantage as speech-speech bilinguals. Emmorey et al. (2008), for example, found that bilinguals who spoke one language and signed another responded comparably to monolinguals on a flanker task of selective attention.

The two main areas of cognitive control where bilinguals are thought to outperform their monolingual counterparts are in selective attention and interference control (see Adesope et al., 2010, for a meta analysis) although these benefits have not always been replicated (Paap et al., 2018; Papageorgiou et al., 2019). The detection and resolution of a linguistic conflict, as would arise in a linguistic violation, is thought to rely on executive functions skills, a set of cognitive skills governing control of a variety of processes (Miyake et al., 2000). The executive functions in question here are selective attention (Wolleb et al., 2017; Ye & Zhou, 2009b; van Herten et al., 2006) and interference control (Ye & Zhou, 2009a; Nee et al., 2007).

2.3.1 Selective Attention

Selective attention, or the capacity to select pertinent information for the resolution of a task while ignoring distractor stimuli, is thought to play a significant role in the ability of bilinguals to control which language they are accessing at a given time. As such, the practice of such a skill would have benefits beyond simply selecting between languages. Bialystok (2001), for example, found that bilingual children have an advantage in selective attention and inhibition which is argued to be associated with enhanced frontal lobe effectiveness. Costa et al. (2008) found that results on the Attention Network Task (ANT; Fan et al., 2002) supported the hypothesis of greater attentional control by bilinguals in the alerting and executive control network. Friesen et al. (2015) found that bilinguals outperformed their monolingual counterparts in response speed during a challenging visual search task suggesting that bilinguals have a better control of visual attention. Bilinguals have also been shown to have an advantage in auditory attention in addition to visual, and in fact this auditory advantage is shared by late bilinguals (Bak et al., 2014). The fact that late bilinguals appear to demonstrate some effects of bilingual advantage may provide insight into neural plasticity and our understanding of second language acquisition and development.

2.3.2 Interference Control

Interference control, or interference suppression, is the capacity to ignore or suppress misleading information which would lead to a faulty response in the task at hand (Bunge et al., 2002). Bilinguals must be good at this in order to suppress a) external linguistic stimuli which are not pertinent to the current language use, and b) internal linguistic information that would lead to a faulty linguistic or behavioral response. In this way, bilingualism is thought to boost frontal lobe functions (Bunge et al., 2002). The Stroop task (Stroop, 1935) is generally considered as the gold standard for linguistic interference, and accordingly is often used to test interference control. Bialystok et al. (2008) found that bilinguals across the lifespan, demonstrate a lower cost in a Stroop task wherein they make fewer errors, and respond quicker than their monolingual counterparts.

It has been proposed that a general inhibitory process is applied to the non-used language in order to avoid interference effects in the selected language (A. M. Philipp et al., 2007; A. Philipp & Koch, 2009). Another proposal is that rather than suppression, or the inhibition of irrelevant stimuli, bilinguals show an advantage in the positive selection of information relevant to the task at hand (Treccani et al., 2009). A meta-analysis conducted by Donnelly et al. (2019) found that bilinguals demonstrate a small but significant advantage on interference control tasks. Interestingly, they found that late bilinguals demonstrated greater advantages than early bilinguals.

2.4 Neural Bases of Control

We now turn to which regions of the brain have been identified as comprising the system of cognitive control which is so crucial for bilinguals. There seems, in fact, to be general agreement in the literature that the anterior cingulate cortex (ACC) and dorsolateral prefrontal Cortex (dlPFC), particularly on the left side, are implicated in this system both in linguistic and non-linguistic tasks. A meta-analysis conducted by Nee et al. (2007) found that the left dlPFC, ACC, and posterior parietal cortex were engaged in interference tasks. Ali et al. (2010) further found that the caudate was activated in order to control interference during a Stroop task.

This system seems to be especially active when bilinguals are asked to switch between their languages, even across different types of tasks. Wang et al. (2007), for example, found increased activation in the dlPFC and ACC when subjects switched into their L2, a finding which is consistent with the idea that bilinguals must inhibit their L1 to speak in their L2. Fan et al. (2003) found that the ACC and left prefrontal cortex show a common effect of linguistic conflict. Hernandez et al. (2000) found more activation in the left dlPFC when participants were switching between naming pictures in English and Spanish. J. Price (1999) further found that language switching or mixing induced increased frontal and parietal activity, consistent with ongoing inhibitory activity to support the selection of a relevant response for the task in the face of competition. In another picture naming task, Abutalebi, Annoni, et al. (2007) found that naming pictures in a dual language condition induced more extensive activation in the left pre-frontal cortex (PFC), the ACC, and the left caudate than did the same task in a single language condition. They further found that this activation was intensified when participants were using their weaker L2, supporting the importance of these regions in selecting a language in the face of interference. The same control system seems to be in use even in late bilinguals, as was found by Crinion et al. (2006) in a semantic decision task. Kovelman, H Shalinsky, et al. (2008) found that the system of control required for speech-speech bilinguals seems to behave slightly differently than for speech-sign bilinguals. In a language switching task, speech-sign bilinguals did not show a significant increase in prefrontal activation when they switched between their two languages. Blanco-Elorrieta et al. (2018), however, found increased activity in the dlPFC and ACC in a MEG study with bimodal bilinguals when they were asked to switch between languages. Control in bimodal bilinguals may behave differently, though, as both languages can be produced simultaneously without significant influence on one another. Inhibition of the non-selected language is therefore less crucial.

2.5 Proficiency in Cognitive Control

As we are investigating processing in non-native speakers, the question of proficiency in the L2 becomes crucial if we wish to consider other variables, especially in neuroelectrophysiological responses. We turn now to what effects proficiency has been shown to have in neurophysiological responses to linguistic tasks. Highly proficient late bilinguals behave neurocogntiviely differently than monolinguals, early bilinguals, and lower proficiency late bilinguals. Kovelman, Baker, & Petitto (2008) showed that despite demonstrating a processing profile that appears similar to native speakers, high proficiency bilinguals show evidence of higher processing costs in the L2, as well as recruitment of control regions.

Abutalebi & Green (2007), however, demonstrated that as proficiency in L2 increases, the relative difference in activation between L1 and L2 decreases. These data point to a decrease in control effort which comes with higher proficiency in the L2. Abutalebi, Brambati, et al. (2007) found that switching into a more dominant language seemed to require less neural resources to suppress the activation of the less dominant language than the other way around.

3 Predictions

Given that contemporary French is -V2, we would predict the ERPs to resemble those of the native English speakers in Andersson et al. (2018), i.e., less native-like than the German learners. If the French-Swedish bilinguals have adapted their parsing strategies to the typological characteristics of Swedish, then a larger posterior P600 should be expected in the word order violation condition in comparison with the control condition during reading of the Swedish sentences (Weyerts et al., 2002). This P600 effect should be modulated by some of the linguistic factors, particularly level of proficiency and level of exposition. Moreover, based on the cognitive control tasks, we predict a modulation of the frontal P300, as well as the frontal P600 components. The predicted effect is that low attentional capacity would result in a less robust P300 component, but a greater amplitude of the frontal P600 because more effort must be exerted to resolve the conflict that was not detected as early.

4 Methods

The methods described herein are designed to replicate those used by Andersson et al. (2018), with the addition of the measurement of selective attention and interference control. The materials are used with the permission of the original authors. All Matlab, Python, and R code used for the presentation, processing, and analysis in this project are original works of the student.

4.1 Participants

The participants collected in this preliminary study were four native French speakers (2 female) living in France who had learned Swedish as a foreign language. All participants were between the ages of 18 and 35 (table 1), and had normal or corrected to normal vision, reported normal hearing, and had no history of neurological or language disorders. All participants filled out a questionnaire about their handedness, socioeconomic status (SES), and language background, including about their age of acquisition (AoA) and length of exposure (LoE) to Swedish.

Note that participant 3 grossly overstated their proficiency in Swedish, and as such will be excluded from further analyses as an outlier (table 1).

The EEG data collected for use as the native Swedish baseline were collected from participants at the Humanities Lab at Lund University in Sweden by Andersson et al. (2018). These were 20 native speakers of Swedish (8 female, mean age: 23;10, SD: 4;9).

4.2 Tasks and Stimuli

This study employs two linguistic tasks–the Acceptability Judgment Task (AJT) and Sentence Completion Task (SCT)–and two tasks targeting executive functioning–Stroop (Stroop, 1935) and Navon (Navon, 1977). All linguistic stimuli deployed (i.e., those in the AJT and SCT) are those already validated in Andersson et al. (2018) and for that reason were not re-piloted for this study.

4.2.1 Sentence Completion Task

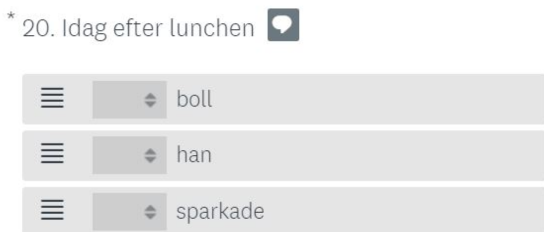
To assess written production of sentences with grammatical word order, a Sentence Completion Task (SCT) was used. Participants were presented with a lead-in fragment on a computer screen followed by boxes with words or short phrases. They were instructed to put these words and phrases in the correct order by ordering them from "1" to "3" so that the sentence is grammatical (fig. 1). Participants were instructed to drag and drop the responses into the right order. In experimental sentences (60), the lead-in fragment consisted of one of two adverbials (”today” or ”at home”), and the boxes contain the subject—third person singular of either a noun (”the girl”), or a personal pronoun (”she”); verb in the simple past; and object of the sentence to be ordered. Half of the sentences had long prefields with additional prepositional modifiers (”Today after lunch”). Experimental sentences were intermingled with fillers (180), consisting of four sentence types: topicalizations (90), questions (30), SVX sentences (30), and negated sentences (30). See table 2 for examples. Sentences were pseudo-randomized such that no more than three sentences from the same condition could appear in series. Stimuli were presented using SurveyMonkey (SurveyMonkey Inc., 2019). Participants were given up to 30 minutes to complete the task.

Participant	Age (yrs)	Sex	AoA (yrs)	LoE (yrs)	Proficiency
1	22.7	F	19	3	7.75
2	23.7	F	22	1.5	7.0
3	20.7	M	15	2	1.0
4	34.1	M	32	1	6.75
Mean:	25.3	N/A	22	1.88	5.63

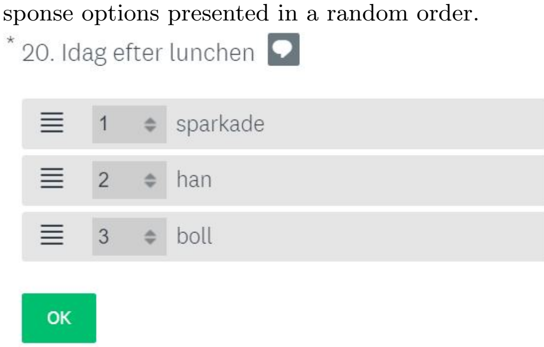
Table 1: Demographic information about native French participants. AoA = Age of Acquisition, LoE = Length of Exposure. Proficiency is performance on the Swedex test, and is scored out of 10.

Sentence Type	**Example**
Experimental (60)	*Hemma tvättade hon filten* Home washed she blanket.DEF ‘At home she washed the blanket’
Topicalization (90)	*Idag på morgonen handlade han alla de nybakade kakorna* Today on morning traded he all the fresh.baked cookies ‘Today in the morning, he sold all the freshly baked cookies’
Question (30)	*När efter lunchen byggde jag ett litet fågelbord?* When after lunch built I one small bird.feeder ‘When after lunch did I build a small bird feeder?’
SVX (30)	*Jag lovade att vara hemma i tid* I promised to be home in time ‘I promised to be home in time’
Negated (30)	*Jag åkte inte med flyg* I went not with flight ‘I did not go by plane’

Table 2: Sentence types used in the AJT and SCT.



(a) Initial sentence presentation in SCT with response options presented in a random order.



(b) Sentence presentation after participant has dragged the words into the correct order. Only then are they presented with the "OK" button to continue to the next item.

Figure 1: Schematic of a trial for the Sentence Completion Task (SCT). In each trial, participants saw a sentence with an initial fragment, followed by several words which would follow, presented in a random order (1a). Participants were instructed to drag and drop the words so that the sentence read grammatically from top to bottom (1b).

4.2.2 Acceptability Judgment Task

To examine how speakers processed sentences offline, we used an Acceptability Judgment Task (AJT). Participants are presented with sentences in the L2 which they must judge as acceptable or unacceptable, indicating their response with a button box. Participants saw 480 (+3 practice) sentences presented word by word on the center of a computer screen in white font on a black background using E-Prime 2.0 software (Psychology Software Tools, Inc., 2016). Each word was presented for 300ms with an inter-stimulus interval of 200 ms. Final words include full stops. The last word is followed by a blank screen for 700 ms, after which three question marks appear until the acceptability judgment is made by pressing a button (either "F" or "J" on a computer keyboard, counterbalanced across lists and participants) corresponding to "good" or "not so good" (fig. 2). Once the judgment is made, participants are presented with a fixation cross for 500 ms, and then the next trial begins.

The 480 sentences are comprised of the types Grammatical V2 (160), Ungrammatical V3 (160), and Fillers (160). To control for wrap-up effects, critical sentences have a final phrase between 0-5 words. There were two lists of stimuli such that sentences having the illegal V3 word order in one list have the legal V2 word order in the other, and vice versa. These lists were counterbalanced across participants.

Sentences were presented in 10 blocks of 48 sentences (about 6 minutes per block), with self-paced pauses between blocks. Blocks were pseudorandomized such that no more than 3 sentences from the same condition (V2 vs V3) would appear in a row, as well as that sentences with the same structure could not appear in close proximity to one another. Participants were presented all sentences from their list exactly once.

The participant was instructed to remain as still as possible while the sentences were on the screen, leaving the index finger of each hand on the response keys on the keyboard. Participants were further instructed not to blink during sentence presentation, but rather to do this while the question marks were on the screen, as we were not interested in the response times for this task.

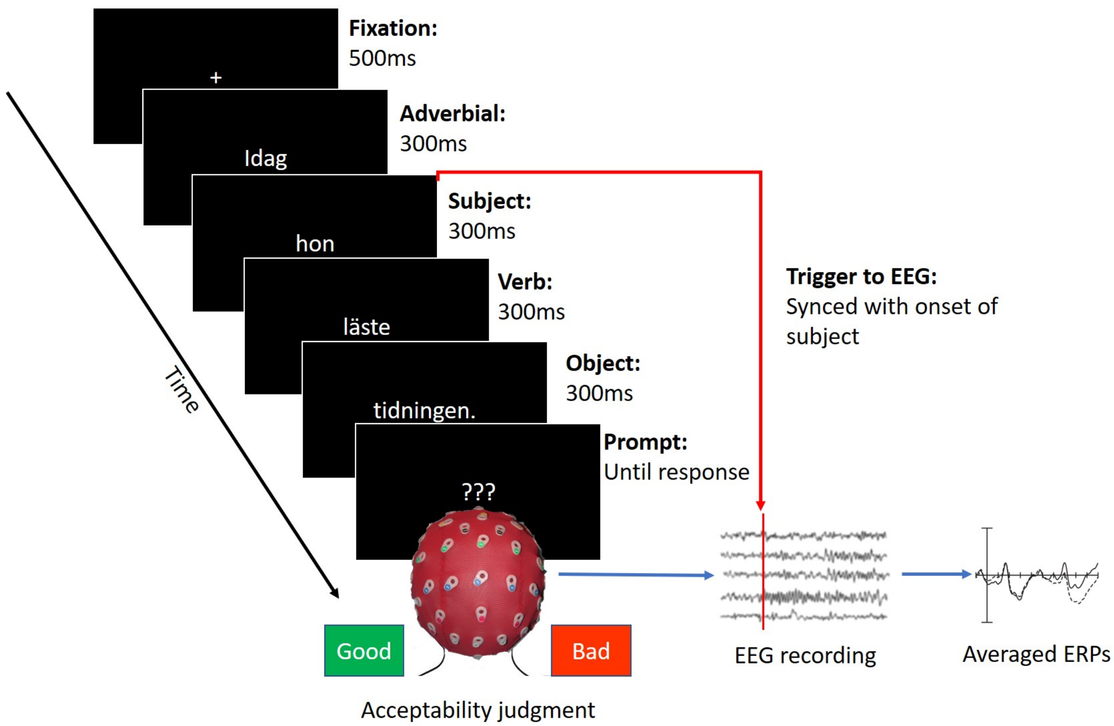


Figure 2: Schematic of a trial for the Acceptability Judgment Task (AJT). Sentences were presented word by word on a screen for 300ms each with 200 ms in between. The last word was presented with a full stop. After the last word, a blank screen would appear for 700 ms, followed by three question marks ("???") until participants made a judgment of "good" or "not so good" via a button press. Following the button press, a blank screen would appear for 1000 ms, followed by a fixation cross in the center of the screen for 500 ms before the next sentence started. EEG triggers were synced with the subject of the sentence.

4.2.3 Executivive Functioning

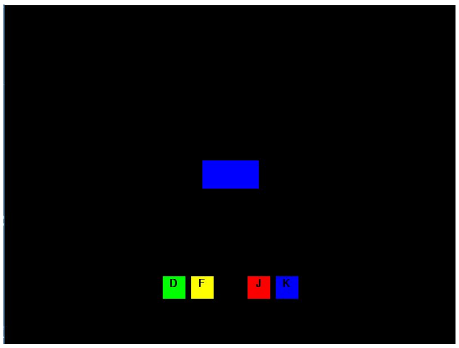
Stroop Task: As a test of interference control, we implemented a Stroop color word task (Stroop, 1935). The design of the task incorporated elements from the Stroop sub-task of the GREFEX battery as well as the Stroop task described in Heidlmayr et al. (2014). The task was divided into several sub-tasks by block, and was administered using ExPyriment, a Python library for cognitive and neuroscientific experiments (Rossum, 1995; Krause & Lindemann, 2014). In each of the blocks, the trial format was the same (fig. 3): participants were presented with a fixation cross in the center of the screen for n ms, where n is a random number between 500 and 1000 drawn for each trial to reduce anticipation effects. A stimulus was then presented in the center of the screen against a black background for a maximum of 1500 ms during which time the participants were asked to respond using one of four keys on a computer keyboard (D, F, J, K) corresponding to that trial's stimulus. These keys were chosen according to natural hand position on a keyboard, so that participants could respond with the index and middle finger on each hand. The color-to-key bindings were randomized for each participant but remained consistent for the whole task for that participant. The keys and colors were displayed on the bottom of the screen from the onset of the fixation cross to the response. If a participant did not respond to a given trial within 1500 ms after stimulus onset, the trial ended. After each trial, a blank screen was presented for 500ms before the next trial began. Each block was preceded by instructions and five practice trials drawn randomly from the block. If participants responded incorrectly to more than two of these trials, they were reminded of the instructions and sent through the practice trials again until they made sufficiently few mistakes to continue onto the experimental trials.

The blocks were organized according to a non-computerized version of this task (Roussel & Godefroy, 2008). In the first block, participants were presented with colored rectangles on a computer screen (fig. 3a). They were asked to respond by pressing one of four buttons corresponding to the color of the rectangle (blue, yellow, green, or red). In the second block, participants were presented with words–*BLEU* (blue), *JAUNE* (yellow), *VERT* (green) or *ROUGE* (red)–and participants were asked to respond according to the *meaning* of the word. These words were presented in white 48pt font (fig. 3b). In the third block, participants were presented with words printed in color and asked to respond according to the *color* of the word. These trials consisted of three types: congruent, incongruent, and neutral. In the congruent trials, participants are presented with a color word printed in the same color, e.g., *VERT* (green) printed in green font (fig. 3c). In the incongruent condition, participants were presented with a color word printed in a different color, e.g., *VERT* printed in red font (fig. 3d). In the neutral trials, participants were presented with non-color words–*CHAT* (cat), *CHIEN* (dog), *MAIN* (hand), or *PIED* (foot)–printed in color. Words were always presented in all capital letters.

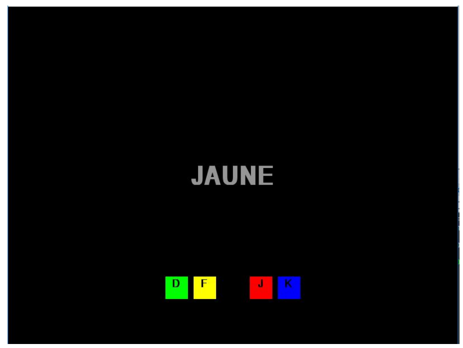
The first two blocks comprised 40 trials (4 colors × 10 scrambles) evenly distributed across the four colors. The third block comprised 128 trials (8 words × 4 colors × 4 scrambles), with a pause after 64 trials. Response times and responses were recorded for comparison across conditions. For purposes of analysis, participants who demonstrated comparably high rates of correctness, and speed of response in incongruent and congruent or neutral trials were considered to have strong interference control skills.

Navon Task: As a test of selective attention, we implemented a Navon task (Navon, 1977). In this task, large letters were presented on the screen, comprised of smaller letters (fig. 4). The large/ global letter could be comprised of smaller/ local versions of itself (congruent – fig. 4a), or of another letter (incongruent – fig. 4b). This task was also presented using ExPyriment (Krause & Lindemann, 2014). Each participant was randomly assigned two target letters from the set {O, S, H, L, A}. Participants were then asked to detect those letters in different conditions by pressing a key on the keyboard, or a different key if they were absent. These keys were counterbalanced across participants.

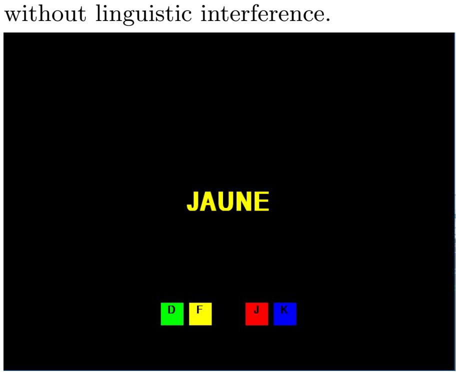
Stimuli were separated into three blocks with different instructions for each block. The trial format was the same for all three blocks. A fixation cross was presented for a period of time ranging from 500 to 1000 ms. The stimulus was presented on the screen for 250 ms, after which the participant had up to 1,500 ms to respond whether they detected their target letters or not. If the participant did not respond within 1,500



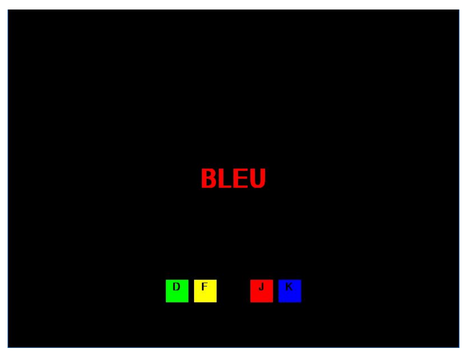
(a) Control trial testing color recognition without linguistic interference.



(b) Control trial testing word recognition without color interference.



(c) Congruent/ facilitation trial where the word and the font color are the same.



(d) Incongruent/ competition trial where the word and font color are different.

Figure 3: Example stimuli for the Stroop Task. To test for interference control, a Stroop task was used with trials in three blocks. In the first block (3a), participants were asked to identify colored rectangles by pressing keys on the keyboard. In the second block (3b), participants were asked to identify color words printed in black and white. In the third block, participants were asked to respond only to the text color of the word printed on the screen. This word might be a color word printed in the same color (congruent – 3c), a color word printed in a different color (incongruent – 3d), or a non-color word printed in color (neutral – not pictured). Participants are assessed on correctness, and speed of response, particularly the difference in response time between congruent/ faciliation trials, and the incongruent/ competition trials, where the participant must inhibit the extraneous linguistic information and respond only to the color.

ms, the trial ended automatically. There was a 500 ms inter-stimulus interval after which the fixation cross was presented as the start of the next trial.

In the first block, participants were asked to indicate if their assigned letters were on the screen either as the global or the local letter. In the second block, participants were asked to indicate only if their assigned letters were on the screen at the global level, and in the third block, only if they were present at the local level.

Each block comprised 50 trials (5 global letters × 5 local letters × 2 scrambles). Participants who demonstrated comparable rates of correctness and speed of response at the global and local level were considered to have strong selective attention skills.

4.2.4 Proficiency and Questionnaires

Swedish proficiency test: As a measure of formal Swedish proficiency, the Word and Grammar section of Swedex (Swedex, 2012) was used, targeting the B1 level of the Common European Framework of Reference for Languages. Participants were required to fill in blanks in a paragraph with the appropriate Swedish word such that the sentence would make sense. Participants used a Swedish keyboard for this task to facilitate writing Swedish letters with diacritics.

English proficiency test: As a measure of English proficiency, the Oxford placement test 2 (Allen, 1992) was used. English proficiency had been voiced as a possible confounding variable, so this task was implemented as a control thereof. In a series of paragraphs, participants selected among three possible choices to fill in blanks, such that the paragraphs would be coherent and grammatical.

Questionnaires: In order to gather demographic information, participants were asked to fill in a questionnaire about language history (Gullberg & Indefrey, 2003), handedness (Oldfield, 1971), and socioeconomic status or SES (Hollingshead, 1975). The questionnaire also included questions about any formal linguistics training participants may have had.

4.3 Procedure

An experimental session took place as follows:

1. Participant signed a consent form presented in their native language (see appendix B), and was assigned to one of two groups, which would determine which list of sentences they would see in the AJT.
2. Participant was fitted with an EEG cap while they filled out a questionnaire regarding their language background, handedness, and socioeconomic status (approximately 15 minutes).
3. Participant completed the AJT during EEG recording (approximately one hour).
4. Participant completed the Swedish proficiency test (approximately 10 minutes).
5. Participant completed the SCT (30 minutes, timed).
6. Participant completed the Stroop task (approximately 10 minutes).
7. Participant completed the Navon task (approximately 10 minutes.
8. Participant completed the English proficiency test (approximately 10 minutes).
9. Participant was thanked, appropriately debriefed, and remunerated EUR 45 for their time.

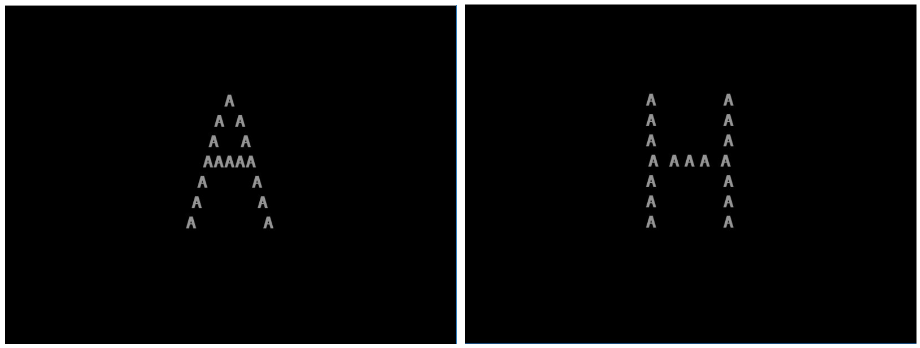
Total session time was between 2.5 and 3 hours depending on how long participants took on the various tasks, as well as how many and how long their breaks were. Participants completed experimental tasks in a dimly lit room.

With the exception of the questionnaire, all tasks were presented on a computer screen at eye-level about 40 centimeters from the participant. A normal AZERTY keyboard was used for all tasks except the Swedish proficiency test.

4.4 EEG recording procedure

Electrophysiological, that is electroencephalographic (EEG) and electrooculographic (EOG), responses were recorded while participants read the sentences in the AJT[1]. Triggers were sent to the EEG system time-locked with the critical word, i.e., the grammatical subject—the point at which the word order violation could first be detected.

[1] Additional details on the EEG recording procedure, as well as the experimental stimuli are available in the online supplementary materials: `https://jeremyyeaton.github.io/supplementaryMaterials/materialsIndex`



(a) Congruent trial where the global (A) and local (A) levels match

(b) Incongruent trial where the global (H) and local (A) levels do not match

Figure 4: Example stimuli for the Navon Task. To test for selective attention, a Navon task was used with trials in three blocks. At the beginning of the experiment, participants were assigned two target letters. In the first block, participants were asked to indicate whether their target letters appeared at either the global or local level by pressing a key, or another key if their letters were not there. The procedure was the same for the second block, except that participants were asked only to search for their target letters at the global level. In the third block, they were asked to search only at the local level.

The EEG data were recorded from 64 active electrodes using a BioSemi ActiveTwo system. Electrodes were mounted in an elastic cap. Electrodes were placed according to the 10-20 electrode positioning system (fig. 12a). Only a subset of these electrodes was used for the ERP analyses discussed here, to match the electrodes used in the original study (fig. 12b). Four additional electrodes were used to record horizontal electro-oculogram (HEOG) and vertical electro-oculogram (VEOG) to monitor eye movements and blinks. These were placed above and below the left eye for VEOG, as well as at the outer canthi of both eyes for HEOG. Two more electrodes were placed on the mastoid process behind each ear to be used in offline re-referencing.

During recording, all scalp electrodes were referenced to the CMS and DRL electrodes which straddle the POz electrode. Data were re-referenced to the average of the two mastoid electrodes during offline processing. The EEG signal was amplified using the ActiveTwo amplifier (bandpass .05 - 100 Hz) and digitized at a sampling rate of 512 Hz.

4.5 Data treatment and analyses

4.5.1 Behavioral analyses

For each of the behavioral tasks (AJT, SCT, Stroop, Navon), responses were analyzed as follows:

Acceptability Judgment Task: During the AJT, participants were instructed to take their time to respond correctly, as well as to use the time while the question marks were on the screen to blink or move as needed. As such, we did not record response times during this task. Only the accuracy of the response was recorded for each trial. Based on this information, a d-prime (d') score (Wickens, 2001) was calculated for each participant using the neuropsychology package in R (R Core Team, 2019; Makowski, 2016). The d' is a signal detection-theoretic measure estimating discrimination between the two conditions based on hits, misses, false alarms, and correct rejections. A score of d' = 0 means that participants could not reliably detect the difference between conditions (or responded at chance), and d' = 4 means that the participant could perfectly discriminate between the two conditions.

Sentence Completion Task: For each sentence in the SCT, accuracy was determined according to whether the participant ordered the parts in the correct V2 word order or in the incorrect V3 word order. This was collapsed into a mean accuracy score by participant.

Stroop Task: For each trial in the Stroop task, accuracy of response and response time (RT) were recorded. Because we are interested here in the interference control of our participants, we included data only from colored word trials in the congruent, incongruent, or neutral condition. We are interested in the relative facilitation and interference effects that can be observed. As such, the difference in mean RT was taken between the congruent and neutral conditions (size of facilitation effect in RT), and the incongruent and neutral conditions (size of interference effect in RT), as well as the overall difference between congruent and incongruent. In addition, the accuracy of each trial was recorded, and the percentage correct was compared between the congruent and neutral, and incongruent and neutral conditions to measure whether there was any facilitation or interference effect in accuracy respectively.

Navon Task: For each trial, the RT and accuracy was recorded. Based on this, each trial was coded as being either a Hit, Miss, False Alarm, or Correct Rejection, and a d-prime score was calculated for the global and local conditions, measuring how well the participant could detect the presence of their target letters as either the large letter or its components, where d' = 4 would mean no mistakes, and d' = 0 would mean effectively random responses. In addition, the mean RT was calculated for hits in both the global and local conditions, as well as the difference between the two.

4.5.2 EEG Data preprocessing

Offline, the data were filtered between 0.5 and 40 Hz. The 40 Hz threshold was selected to reduce high-frequency noise, including electrical line noise. The 0.5 Hz threshold was used to reduce slow drift effects. The three bipolar channels (HEOG, VEOG, and Mastoid), were meaned as a single channel each. The data then underwent an initial automatic artifact rejection routine aimed at eliminating anomalous segments of data which fell outside of a given z-score threshold (4 for EOG channels, 20 for EEG channels), presented a range larger than 1500 μV, or had sudden, extreme changes in amplitude. Once automatic rejection took place, data underwent visual inspection to remove any further highly anomalous segments or channels before passing the data to Independent Component Analysis (ICA). ICA was implemented using "runica" from EEGLAB (Delorme & Makeig, 2004) for 25 components. Based on the ICA, ocular, movement, and electrode artifacts were identified and removed based on their scalp topographies, and component time series[2]. The data then underwent visual inspection to remove any residual ocular artifacts. After artifact rejection was complete, channels that had previously been removed due to noise or artifacts were interpolated based on an average from their neighboring electrodes. This step was conducted after the ICA decomposition so that all channels used in ICA would be independent. Time-locked ERPs were excised from the continuous EEG signal offline for each participant at each electrode site in 1100 ms epochs, using a 100 ms pre-stimulus baseline. Only trials for which a correct response was supplied during the AJT were included in subsequent analyses. To be included in further analyses, a participant must have had at least 10 artifact-free trials per condition.

ERP processing was conducted using FieldTrip (Oostenveld et al., 2011). EEGLab (Delorme & Makeig, 2004) was used to read in the raw data, as well as for its ICA functionality. Both of these are toolboxes for Matlab (MATLAB, 2017).

Because the baseline native Swedish data was recorded on a different system than we used here, it was recorded at 500 Hz, not 512. In order to remedy this sampling rate mismatch, all data were downsampled to 256 Hz so that they could be compared on the same axes.

4.5.3 ERP Analyses

Once ERPs were segmented into 1100 ms epochs, mean amplitude was taken in each of the following time windows of interest: 300-500, 500-700, 700-900, and 900-1000 ms post stimulus. These time windows were chosen by Andersson et al. (2018) based on earlier studies targeting word order violations as well as from their inspection of individual waveforms. We will take the mean amplitude also in the 200-400 ms window. This is targeted at the P300 component that is crucial to our investigation of attentional processes. As a post-hoc measure, we also decided to investigate the 0-200 ms window as well.

[2]For more information on the process of cleaning data of ocular artifacts, see fig. 13.

4.5.4 Planned statistical tests

It would be imprudent to perform parametric statistical tests on this sample because it is too small[3]for these methods to be useful. As such, we cannot here use the same statistical testing methods as Andersson et al. (2018). As such, we will present mostly descriptive statistics on the EEG data at this time.

As an exploratory measure, t-tests were conducted in the time windows of interest to investigate whether the difference in amplitude between the two conditions was sufficient to reject the null hypothesis.

Upon further data collection, the mean amplitudes defined above will be subjected to a repeated-measures ANOVA with Word order (V2/ V3), Hemisphere (right/ left), Lateral position (lateral/ medial), and Anterior/ Posterior position or Ant/ Post (frontal/ fronto-temporal/ temporal/ central/ parietal/ occipital) as the four within-subjects factors per Andersson et al. (2018).

We also plan to conduct a Linear Mixed Model to investigate the relationship between our offline measures of proficiency and cognitive functioning and the amplitude of the ERP during the time windows in question. For this, amplitude of EEG is the dependent variable, the fixed effects are word order (V2/ V3), AJT d', SCT accuracy, proficiency score, Stroop effect size, Navon effect size, and electrode location. The random effect is participant.

For exploratory reasons, we have also looked here at whether ERP amplitude is predictive of Swedish language performance, and/or accuracy or response time on the cognitive tasks. This has been done by using simple Pearson's correlations between ERP amplitude and our offline measures in the windows of interest and sites that are relevant to our hypotheses. We set $\alpha = 0.01$ to reduce noise from multiple comparisons. The statistical power here is far too low to draw conclusions, but this was done to see whether there were trends in the data consistent with our hypotheses. At the moment, all of the RT measures from the Stroop and Navon task are presented without transformations. A transformation such as a log transform or taking $\frac{1}{RT}$ (speed) might be used in the future to achieve a normal distribution for statistical testing.

[3]For a variety of logistical and administrative reasons, we were unable to collect as many bilingual participants as we would have liked prior to the time of writing. We hope to collect more participants prior to the oral defense.

5 Results

Because participant 3 was an outlier both in terms of behavior and EEG, they have been excluded from the results presented here.

5.1 Linguistic Behavioral Results

The behavioral results on the Swedish language tasks (AJT, SCT, Swedex) are somewhat internally inconsistent (table 3). While all three participants performed quite well on the Swedex proficiency test (all above 6.5/10), this measure of proficiency was not strictly predictive of their performance on the other two tasks. In particular, participant # 2 performed quite well on the Swedex task, but near chance on both the SCT and AJT. The other two participants could both reliably detect (AJT) and reliably produce (SCT) the correct V2 word order in Swedish. We thus here present a high, medium, and low proficiency Swedish speaker in this sample.

Of note, however, is that the linguistic measures do not seem to correlate with one another such that performance on the Swedex does not provide much of a prediction of performance on the AJT or SCT. The sample presented here is very small so it is impossible to say anything definitive about this but it is something that should be kept in mind moving forward.

5.2 Cognitive Control

Participants performed at or near ceiling in accuracy on the Stroop and Navon task. As such, no effect can be observed in either accuracy (Stroop) or d' (Navon) between the experimental conditions (table 4). We do, however, see differences in response time (table 5).

5.2.1 Stroop

If a Stroop effect were observed in the accuracy domain, then we should see higher rates of correctness in the congruent condition than the neutral condition (facilitation), and the reverse for the incongruent condition (interference). Because the participants performed at ceiling in terms of accuracy across all conditions, this effect is not observed.

Participant	AJT (d’)	SCT (%)	Proficiency (SW)	Proficiency (EN)
1	4.32	93.1	7.75	84.31
2	0.03	37.8	7.0	66.67
4	3.55	96.1	6.75	72.55

Table 3: Behavioral scores on Swedish language tasks. For the AJT, d-prime scores are reported where d’ = 0 indicates chance performance and d’ = 4 indicates near-perfect discrimination between the conditions. For the SCT, percentage correct is reported. Proficiency scores (out of 10) from the Swedex test–Proficiency (SW)–are reproduced here for ease of reference. We also report here scores on the English proficiency task–Proficiency (EN)–as a percentage of correct responses.

Participant	Facilitation	Interference	Global d’	Local d’
1	0.00	-0.08	3.54	3.91
2	0.03	0.03	3.54	3.34
4	0.03	0.03	3.91	3.90

Table 4: Accuracy measures on cognitive tasks. Facilitation and Interference are accuracy in the neutral condition minus accuracy in the congruent and incongruent conditions of the Stroop task respectively. Because participants performed at ceiling across conditions, no effect is observed. Global d’ and Local d’ are from the two conditions of the Navon task. Again, due to high accuracy across both conditions, no effect is observed.

Participant	Interference Cost	Facilitation Gain	Difference (Stroop)	Global RT	Local RT	Difference (Navon)
1	-28.36	47.84	76.21	368.23	460.35	92.12
2	8.54	28.13	19.58	309.36	398.80	89.44
4	-24.93	-31.09	-6.17	232.35	268.89	36.54

Table 5: Response time (RT) measures on cognitive tasks. Interference cost and Facilitation gain are the response time in the neutral condition minus the response time in the incongruent and congruent conditions of the Stroop task respectively. Negative numbers mean that the response time was slower in that condition than in the neutral one. Difference (Stroop) is calculated as the RT in the incongruent condition minus the RT in the congruent condition. Negative numbers mean that the RT in the incongruent condition was faster. Global RT and Local RT are mean RTs for hits in each of those conditions in the Navon task. Difference (Navon) is calculated as Local RT minus Global RT. This difference is significant for all subjects taken together ($p = 0.005$), but not for all subjects taken individually.

In the RT domain, we should see an increase in RT in the incongruent condition relative to the neutral condition demonstrating interference, and a reduction in RT in the congruent condition relative to the neutral condition demonstrating facilitation. The Stroop effect is bidirectional in this regard, accounting both for the positive and negative influence of linguistic information on performance. Put differently, if an individual were not at all influenced by the linguistic content of a color word presented, then there should be no difference in response time relative to the neutral condition in either the positive or the negative. As such, the larger the difference between the conditions, the more susceptible that individual is to linguistic interference.

We see evidence of this effect in the RT data collected. Two out of three participants responded slower in the incongruent condition than the congruent one. Unfortunately it seems that the number of trials was too low to reliably capture the effect in a significant way, but we see a trend nonetheless, pointing to some small effect of facilitation and interference.

5.2.2 Navon

If a Navon effect were observed in the accuracy domain, we should see lower rates of detection in the Local condition than the Global one. This does not seem to be borne out by the data collected, with one participant even appearing to show the opposite effect having a lower d' score in the Global condition than the Local one. This is not strictly bad *per se*, as the Navon effect seeks to measure selective attention. If participants can selectively attend to the appropriate level of information equally in both conditions, then they have high attentional control.

In the RT domain, the Navon effect would be demonstrated via higher RTs in the Local condition than the Global one, indicating increased cost of detection. We do see this difference borne out by the data, with all three participants responding slower in the Local condition. This increased cost of detection is significant for all the participants together ($p = 0.005$, two-sample t-test), but only for some participants taken individually. For our purposes here, the smaller the difference between the two conditions, the better attentional control demonstrated by that participant, whereas a larger difference indicates poorer selective attentional skills.

5.3 ERP Results

Unfortunately, because the original Swedish data in Andersson et al. (2018) were recorded on a different type of system than we used, the data were not comparable in their raw form. The Swedish monolingual and French bilingual data will therefore be presented separately on different scales.

In all of the ERP plots presented here, the x-axis is time in seconds, where $t = 0$ is stimulus onset, and the y-axis is EEG amplitude in μV. Note that ERPs presented here are plotted with positive upwards[4]. Overall, however, the shape, distribution, and latency of the ERPs in the French bilinguals closely resembles those of the Swedish natives (fig. 5). Globally, we see more event-related activity at frontal sites, with a clear bimodal shape with positive peaks around 300 and 600 ms.

If we look at the ERPs by condition on an electrode-by-electrode basis (fig. 6 & 7), we note that the same general shape can be observed for both conditions across most electrodes for both groups, and certainly all of the anterior sites[5] (frontal, fronto-temporal, and fronto-central electrodes; fig. 6). We see the greatest positive amplitude at medial sites (F3, F4, FC3, FC4)[6].

In the native Swedish speakers, we see this same bimodal positivity for midline posterior electrodes (fig. 7), but with a somewhat reduced amplitude. In the bilinguals, however, the bimodal positivity is significantly reduced, and almost entirely disappears at medial parietal sites. Over lateral sites, particularly on the left, we see a bimodal negativity instead, with the first peak occurring around 150 - 250 ms,

[4]Although it is a common practice in the literature to plot ERPs with negative upward, I have elected not to do so here for two main reasons. The first is that the components of interest in this study are positivities, it is thus clearer to plot them with positive upwards. The second reason is that to the best of my knowledge the rest of the (cognitive) scientific community, plot positive upwards. As such the plotting of positive upward should make this paper slightly easier to read for those who are not accustomed to looking at ERPs.

[5]For the purposes of description, anterior will refer to electrodes on the nasion aspect of the coronal midline, while posterior will refer to electrodes on the inion aspect. Medial and lateral refer to distance from the saggital midline.

[6]See figure 12 for more information on how the electrodes are identified and located.

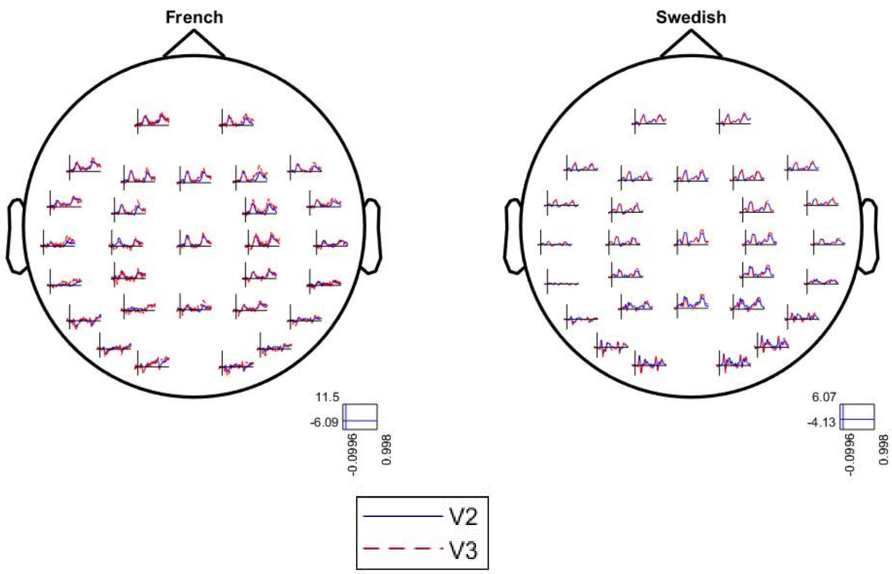


Figure 5: ERP results in French bilinguals and Swedish natives (mean amplitude over time). The vertical line on each plot represents stimulus onset. Note that we see much more activity on frontal sites than posterior ones. Note also that positive is plotted upward and that the y-axis scale is not the same for French bilinguals and Swedish natives.

slightly earlier than the first positive peak at other sites.

Globally, though, we see that the ERP assumes more or less the same shape in both conditions, with a few variations. To examine the differences further, we took the difference between the meaned ERPs (V3-V2), or rather, mean ERP in the V3 violation condition minus mean ERP in the V2 canonical condition. This difference therefore represents the reaction to the error relative to a correct sentence baseline. Thus, if the participants did not have an online response to the error, this line should be flat.

In the Swedish group, we note a significant positivity around 600-800 ms latency (consistent with the P600 component) at anterior, central, and posterior sites, primarily at medial electrodes (fig. 8 & 9). Thus, using the same data, but an independent and slightly different data-processing pipeline, we have reproduced the main effect observed in Andersson et al. (2018). This positivity is greatest at posterior sites, consistent with previous studies eliciting the P600 in response to word-order violations. We also observe a significant negative peak mostly over posterior sites, around 200 to 600 ms, consistent with the N400 ERP component posited to signal linguistic integration processes (Kutas & Federmeier, 2011).

In the French group, we observe the expected P600 component, thus reproducing the effect documented in Andersson et al. (2018) in a different population. This component is strongest over medial anterior sites. While we do observe a more posterior P600 component as well, it is not as pronounced as in the Swedish data, consistent with the English results in Andersson et al. (2018). We also see quite a pronounced negativity with a peak around 150 ms appearing more strongly on the left side–consistent with the ELAN component Steinhauer & Drury (2012)–but also appearing on more posterior sites as well. This early negativity was also observed in the English learners.

Andersson et al. (2018) noted that the difference amplitude over frontal sites was significant across groups in their data. As such, we selected the same electrode subset for further examination (fig. 10). At these electrodes, we find a significant difference between the two conditions in the 300-500, 500-700, and 700-900 ms windows for the French participants ($\alpha = 0.05$). In the Swedish population, however, we find a significant difference only in the 700-900 ms window. Admittedly, the statistical power of this test in the French group is quite low, but we are confident that the same result would be borne out

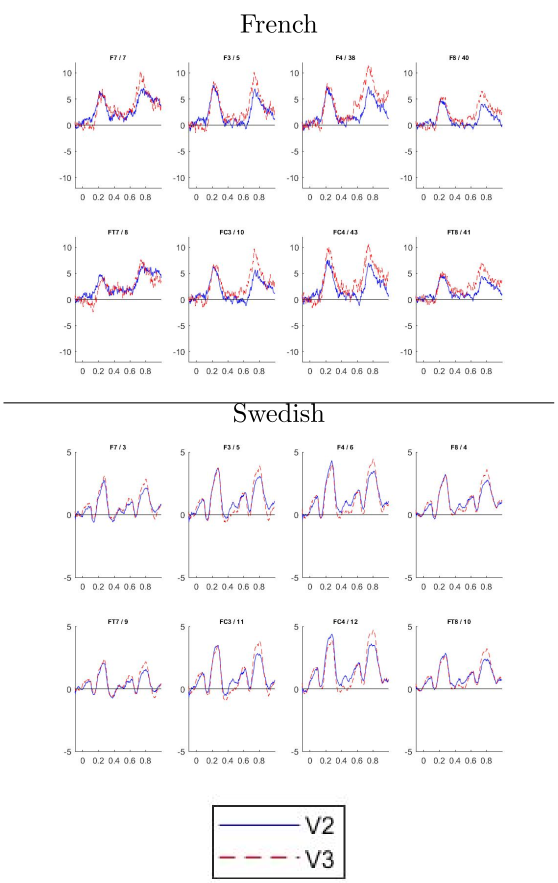


Figure 6: ERPs at anterior sites. We observe a consistent two-peak shape across both language groups and experimental conditions. These peaks are larger at medial sites. Note that the y-axis scale is not the same for French bilinguals and Swedish natives.

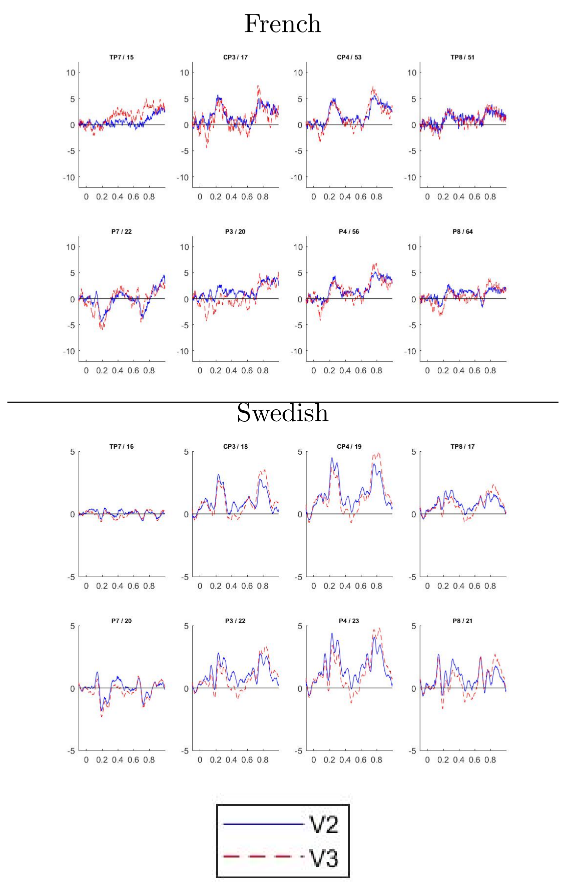


Figure 7: ERPs at posterior sites. A two-peaked ERP can be observed at most medial electrodes.At a left-lateral posterior site, however, a two negative peak shape is observed in both language groups. Note that the y-axis scale is not the same for French bilinguals and Swedish natives.

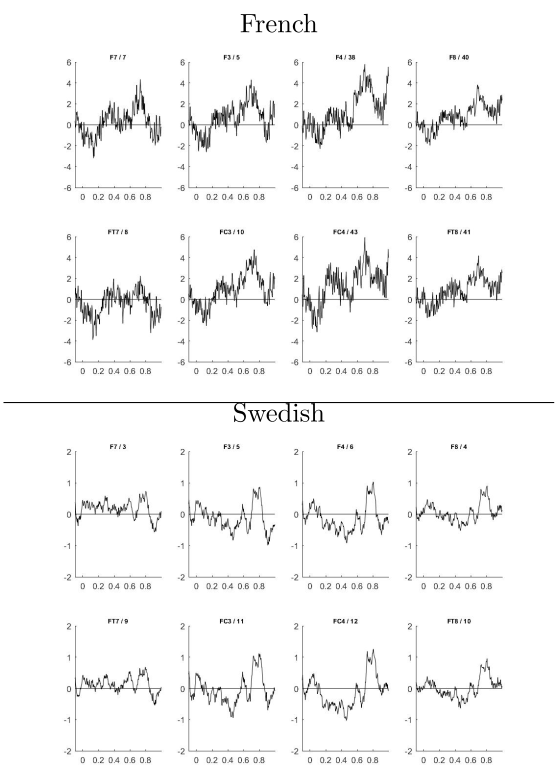


Figure 8: Difference at anterior sites (V3 - V2). We see a clear P600 component in both the learner and native groups. We also note an early negativity in the learner group, which appears to be left-lateralized, consistent with the ELAN component. Note that the y-axis scale is not the same for French bilinguals and Swedish natives.

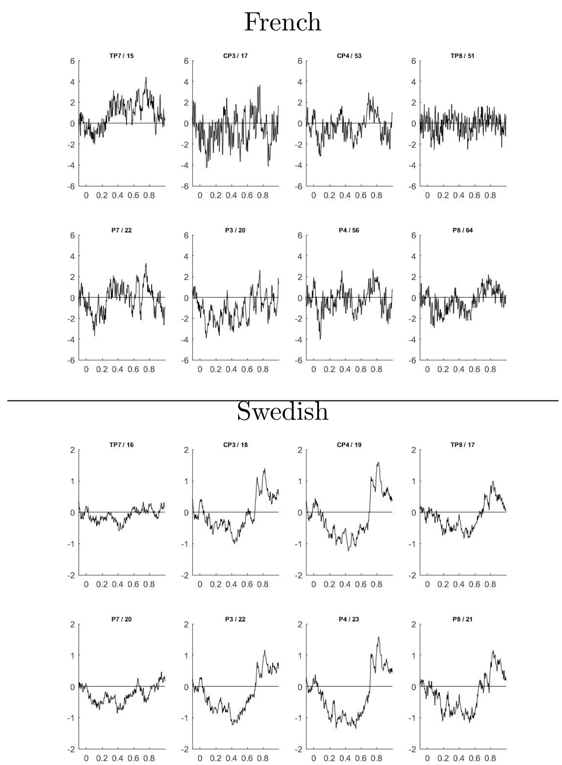


Figure 9: Difference at posterior sites (V3 - V2). A posterior P600 component is observed in the native Swedish group. We see a much weaker posterior P600 in the learner group. Furthermore, we observe a pronounced N400 component (negative peak around 400 ms) over medial posterior sites in the native Swedish group, but not in the learner group. Note that the y-axis scale is not the same for French bilinguals and Swedish natives.

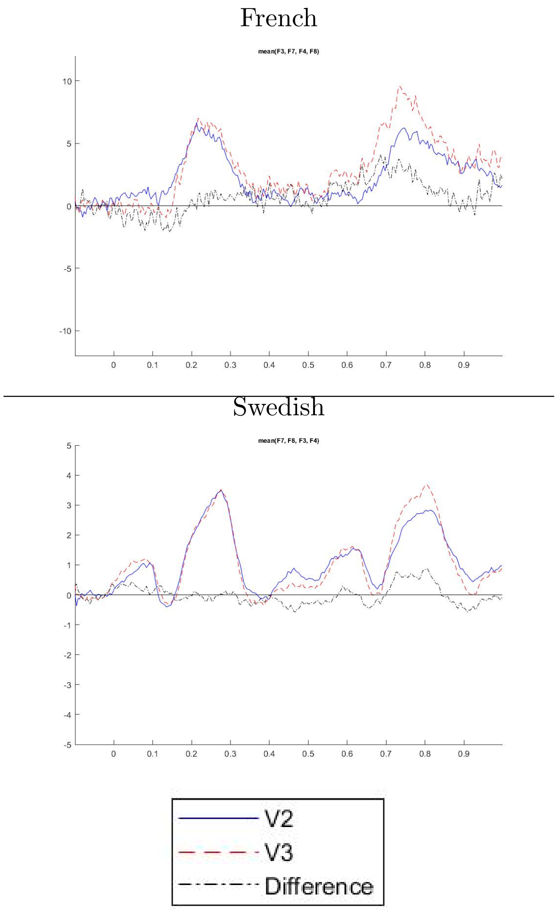


Figure 10: Mean ERP at frontal sites. ERPs meaned over frontal electrodes (F3, F4, F7, F8), where anterior P600 component is strongest across groups. Note that the y-axis scale is not the same for French bilinguals and Swedish natives.

at least in the 500-700 and 700-900 ms windows given additional data.

If we look as well at the topographical distribution of these responses (fig. 11), we see some marked group differences. Most importantly, we see that the scalp distribution of the positivity in the 700-900 ms window is quite different in the Swedish and French groups. Whereas in the Swedish group this positivity is quite central, in the French group we see it is quite frontal. It is possible, however, that a comparable central P600 is occurring in the French group, yet it is simply overshadowed by the high frontal activity, so it is not captured by the scaling. Temporally, we note that the central positivity seems to begin earlier in the Swedish group than the French one, with an effect apparent in the 500-700 ms window. In both groups, we also see a lingering frontal positivity, although this seems to be much more left-lateralized in the learner group than in the natives.

5.3.1 ERPs and Linguistic Behavior

We turn now to the relationship between the online measures (ERPs) and our offline linguistic measures (AJT, SCT, Swedex). We found that the amplitude of the difference ERP negatively correlates with the performance on the SCT in the 0-200 and 500-700 ms windows over left centro-parietal sites. This is consistent with an increased early left negativity (0-200 ms), as well as perhaps a reduced amplitude of the P600 component in higher proficiency speakers. The d' score on the AJT correlated with occipital sites during the 500-700 ms window in the violation condition, and with medial fronto-central sites in the 700-900 ms window in the difference ERP. Thus, higher discriminability seems to point to a more pronounced posterior P600 component and a prolonged anterior positivity in response to errors. No trends were found with the Swedex scores.

5.3.2 ERPs and Cogntive Control

Finally we turn to the relationships between our offline cognitive tasks and the ERPs. Only the results of those tests related to our hypotheses which appear to show significance are presented.

There seems to be a trend between the d' for the Local condition of the Navon task, and amplitude at frontal and parietal medial sites during the 0-200 and 200-400 ms windows. This trend is such that higher discriminability in the challenging attentional condition means greater negativity over these sites during this window. The difference in RT between the Global and Local conditions correlates negatively with central and fronto-central sites during the 500-700 ms window.

Response time on the Stroop task correlated with amplitude at occipital sites in the 500-700 ms window, as well as with frontal and temporal sites in the 200-400 ms window.

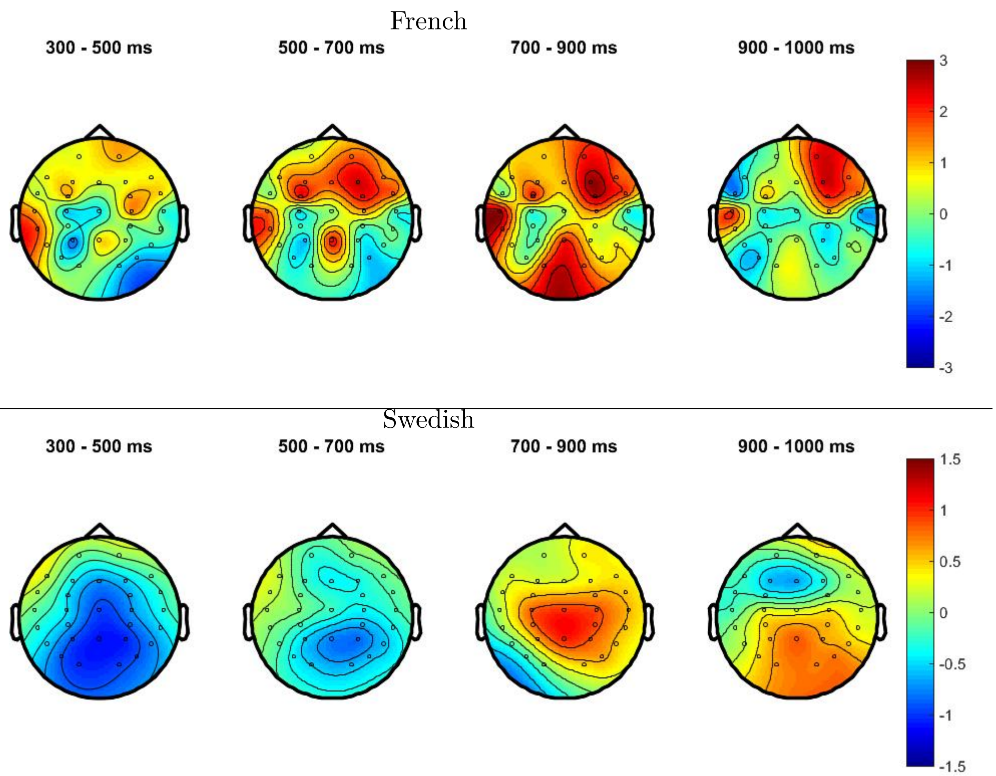


Figure 11: Topographies of V3-V2 difference in time windows of interest. A central/posterior P600 is observed in both groups, however, much more frontal positivity is observed throughout the course of the epoch in the learner group. In the native group, a posterior N400 component is observed, which does not appear in the learner group. Note that the color scale is not the same for French bilinguals and Swedish natives.

6 Discussion

We return now to our predictions. Our first prediction is that the data collected here from participants whose L1 is -V2 would appear more like the other -V2 data collected in Andersson et al. (2018) than the natives or +V2 learners. This seems overall to be borne out by our data. We predicted that the posterior P600 would be greatly reduced in the learner group which appears to be the case as well. The anterior P600 effect, however, was predicted to be more pronounced in our learner group as part of a pronounced positivity over the whole trial epoch. We see this in our data as well. Differences in recording equipment present challenges for comparing these data exactly but perhaps z-score transformations may be useful in the future.

Our other predictions pertain to the relationships between offline linguistic and cognitive measures and ERP amplitude. We predicted that the posterior P600 would be reduced as a function of Swedish language ability. We did not find relationships between Swedex proficiency scores and amplitude for any time windows or regions of interest. It is unclear whether this measure is sufficient to capture proficiency in the language as it does not seem to have any relationship to performance on the other two linguistic tasks which theoretically rely on the same competences. We found, however, that the d' score on the AJT correlated with the amplitude at posterior sites, pointing to a relationship between behavioral performance on the task and online responses. Overall, though, there were no sufficiently robust relationships in our data to support this prediction.

With regard to the relationship between offline measures of selective attention and interference control, we predicted that improved selective attention and interference control would coincide with a more robust frontal P300 component but a less robust frontal P600 component. We find some evidence of the relationship between interference control and the frontal P300 amplitude but our data show a trend in the opposite direction with regard to the frontal P600 effect. Thus, while there are some promising trends, some of our data does appear contradictory to our prediction.

Bear in mind, however, that the sample size presented here is quite small, and lacks the statistical power to definitively support or reject our hypotheses. Upon the collection of further data, we hope to conduct more rigorous statistical tests and models investigating the relationship between the offline and online measures.

7 Conclusions

We have documented here the main effect found in Andersson et al. (2018) with a different population. Thus, the French results here appear to show the same shape as the results of the English learner group which is also -V2, but not the German learner group which is +V2. Pending further statistical tests, this lends credence to a theory of cross-linguistic influence in L2 acquisition where the presence or absence of V2 word order in the L1 impacts processing in the L2.

We have also shown some evidence of a relationship between cognitive measures thought to play a role in language processing and online responses to linguistic stimuli. In the future, we hope to use additional analyses to determine whether parieto-central activity in the learner group is being outscaled by frontal activity, making the parieto-central activity harder to detect. In the same way, we hope that source localization analyses may help us to disentangle the generators of the frontal P300, and both anterior and posterior P600, providing additional insight into attentional load or structural complexity (Kaan & Swaab, 2003).

In addition to the planned source reconstruction analyses to differentiate the various components, we believe that doing some data-driven cluster or classification analyses on the larger data set may provide additional insight into the linguistic and neurocognitive mechanisms underlying L2 processing.

References

Abutalebi, J., Annoni, J.-M., Zimine, I., Pegna, A. J., Seghier, M. L., Lee-Jahnke, H., ... Khateb, A. (2007, 10). Language Control and Lexical Competition in Bilinguals: An Event-Related fMRI Study. *Cerebral Cortex*, *18*(7), 1496-1505. Retrieved from https://doi.org/10.1093/cercor/bhm182 doi: 10.1093/cercor/bhm182

Abutalebi, J., Brambati, S. M., Annoni, J.-M., Moro, A., Cappa, S. F., & Perani, D. (2007). The neural cost of the auditory perception of language switches: An event-related functional magnetic resonance imaging study in bilinguals. *Journal of Neuroscience*, *27*(50), 13762–13769. Retrieved from http://www.jneurosci.org/content/27/50/13762 doi: 10.1523/JNEUROSCI.3294-07.2007

Abutalebi, J., & Green, D. (2007). Bilingual language production: The neurocognition of language representation and control. *Journal of Neurolinguistics*, *20*(3), 242 - 275. Retrieved from http://www.sciencedirect.com/science/article/pii/S0911604406000716 doi: https://doi.org/10.1016/j.jneuroling.2006.10.003

Adesope, O. O., Lavin, T., Thompson, T., & Ungerleider, C. (2010). A systematic review and meta-analysis of the cognitive correlates of bilingualism. *Review of Educational Research*, *80*(2), 207-245. Retrieved from https://doi.org/10.3102/0034654310368803 doi: 10.3102/0034654310368803

Ali, N., Green, D. W., Kherif, F., Devlin, J. T., & Price, C. J. (2010). The role of the left head of caudate in suppressing irrelevant words. *Journal of Cognitive Neuroscience*, *22*(10), 2369-2386. Retrieved from https://doi.org/10.1162/jocn.2009.21352 (PMID: 19803688) doi: 10.1162/jocn.2009.21352

Allen, D. (1992). *Oxford placement test.* Oxford: Oxford University Press.

Andersson, A., Sayehli, S., & Gullberg, M. (2018, 7 6). Language background affects word order processing in a second language online but not offline. *Bilingualism: Language and Cognition*. doi: 10.1017/S1366728918000573

Bak, T. H., Vega-Mendoza, M., & Sorace, A. (2014). Never too late? an advantage on tests of auditory attention extends to late bilinguals. *Frontiers in Psychology*, *5*. doi: 10.3389/fpsyg.2014.00485

Bialystok, E. (2001, 01). Bilingualism in development: Language, literacy, & cognition. *Bilingualism in Development: Language, Literacy, and Cognition.*. doi: 10.1017/CBO9780511605963

Bialystok, E., Craik, F., & Luk, G. (2008, 08). Cognitive control and lexical access in younger and older bilinguals. *Journal of experimental psychology. Learning, memory, and cognition*, *34*, 859-73. doi: 10.1037/0278-7393.34.4.859

Bialystok, E., Craik, F. I., Green, D. W., & Gollan, T. H. (2009). Bilingual minds. *Psychological Science in the Public Interest*, *10*(3), 89-129. Retrieved from https://doi.org/10.1177/1529100610387084 (PMID: 26168404) doi: 10.1177/1529100610387084

Blanco-Elorrieta, E., Emmorey, K., & Pylkkänen, L. (2018). Language switching decomposed through meg and evidence from bimodal bilinguals. *Proceedings of the National Academy of Sciences*, *115*(39), 9708–9713. Retrieved from https://www.pnas.org/content/115/39/9708 doi: 10.1073/pnas.1809779115

Bohnacker, U. (2006). When swedes begin to learn german: from v2 to v2. *Second Language Research*, *22*(4), 443-486. Retrieved from https://doi.org/10.1191/0267658306sr275oa doi: 10.1191/0267658306sr275oa

Bunge, S. A., Hazeltine, E., Scanlon, M. D., Rosen, A. C., & Gabrieli, J. D. E. (2002). Dissociable contributions of prefrontal and parietal cortices to response selection. *NeuroImage*, *17*, 1562-1571.

Colzato, L. S., Bajo, M. T., van den Wildenberg, W., Paolieri, D., Nieuwenhuis, W., S.and La Heij, & Hommel, B. (2008). How does bilingualism improve executive control? a comparison of active and reactive inhibition mechanisms. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 302-312.

Costa, A., Hernández, M., & Sebastián-Gallés, N. (2008). Bilingualism aids conflict resolution: Evidence from the ant task. *Cognition*, *106*(1), 59 - 86. Retrieved from http://www.sciencedirect.com/science/article/pii/S001002770600268X doi: https://doi.org/10.1016/j.cognition.2006.12.013

Crinion, J., Turner, R., Grogan, A., Hanakawa, T., Noppeney, U., Devlin, J. T., ... Price, C. J. (2006). Language control in the bilingual brain. *Science*, *312*(5779), 1537–1540. Retrieved from https://science.sciencemag.org/content/312/5779/1537 doi: 10.1126/science.1127761

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

Donnelly, S., Brooks, P., & Homer, B. (2019, 02). Is there a bilingual advantage on interference-control tasks? a multiverse meta-analysis of global reaction time and interference cost. *Psychonomic Bulletin & Review*. doi: 10.3758/s13423-019-01567-z

Dryer, M. (2005, 01). Order of subject, object, and verb. *The World Atlas of Language Structures*.

Emmorey, K., Luk, G., Pyers, J. E., & Bialystok, E. (2008). The source of enhanced cognitive control in bilinguals: Evidence from bimodal bilinguals. *Psychological Science*, *19*(12), 1201-1206. Retrieved from https://doi.org/10.1111/j.1467-9280.2008.02224.x (PMID: 19121123) doi: 10.1111/j.1467-9280.2008.02224.x

Engel, U. (1974). Syntaktische besonderheiten der deutschen alltagssprache. In H. Moser (Ed.), (pp. 199 – 228). Düsseldorf: Schwann. Retrieved from http://nbn-resolving.de/urn:nbn:de:bsz:mh39-5993

Fan, J., Flombaum, J. I., McCandliss, B. D., Thomas, K. M., & Posner, M. I. (2003). Cognitive and brain consequences of conflict. *NeuroImage*, *18*(1), 42 - 57. Retrieved from http://www.sciencedirect.com/science/article/pii/S1053811902913194 doi: https://doi.org/10.1006/nimg.2002.1319

Fan, J., McCandliss, B. D., Sommer, T., Raz, A., & Posner, M. I. (2002). Testing the efficiency and independence of attentional networks. *Journal of Cognitive Neuroscience*, *14*(3), 340-347. Retrieved from https://doi.org/10.1162/089892902317361886 doi: 10.1162/089892902317361886

Friederici, A., Gunter, T., Hahne, A., & Mauth, K. (2004, 02). The relative timing of syntactic and semantic processes in sentence comprehension. *Neuroreport*, *15*, 165-9. doi: 10.1097/00001756-200401190-00032

Friederici, A., Hahne, A., & Mecklinger, A. (1996, 10). Temporal structure of syntactic parsing: Early and late event-related brain potential effects. *Journal of experimental psychology. Learning, memory, and cognition*, *22*, 1219-48. doi: 10.1037/0278-7393.22.5.1219

Friederici, A., & Meyer, M. (2004). The brain knows the difference: two types of grammatical violations. *Brain Research*, *1000*(1), 72 - 77. Retrieved from http://www.sciencedirect.com/science/article/pii/S0006899303040563 (Brain Research Volume 1000) doi: https://doi.org/10.1016/j.brainres.2003.10.057

Friederici, A., 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. Retrieved from http://www.sciencedirect.com/science/article/pii/0926641093900262 doi: https://doi.org/10.1016/0926-6410(93)90026-2

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, *6*(2), 78 - 84. Retrieved from http://www.sciencedirect.com/science/article/pii/S1364661300018398 doi: https://doi.org/10.1016/S1364-6613(00)01839-8

Friesen, D. C., Latman, V. V., Calvo, A. P., & Bialystok, E. (2015). Attention during visual search: The benefit of bilingualism. *The international journal of bilingualism : cross-disciplinary, cross-linguistic studies of language behavior*, *19 6*, 693-702.

Greenberg, J. H. (1963). Some universals of grammar with particular reference to the order of meaningful elements. In J. H. Greenberg (Ed.), *Universals of human language* (p. 73-113). Cambridge, Mass: MIT Press.

Gullberg, M., & Indefrey, P. (2003). *Language background questionnaire.* Nijmegen: Max Planck Institute for Psycholinguistics.

Hagoort, P. (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. Retrieved from `https://doi.org/10.1162/089892903322370807` doi: 10.1162/089892903322370807

Hahne, A., & Friederici, A. (1999, 04). Electrophysiological evidence for two steps in syntactic analysis: Early automatic and late controlled processes. *Journal of cognitive neuroscience*, *11*, 194-205. doi: 10.1162/089892999563328

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. Retrieved from `http://www.sciencedirect.com/science/article/pii/S0926641001001276` doi: https://doi.org/10.1016/S0926-6410(01)00127-6

Heidlmayr, K., Moutier, S., Hemforth, B., Courtin, C., Tanzmeister, R., & Isel, F. (2014). Successive bilingualism and executive functions: The effect of second language use on inhibitory control in a behavioural stroop colour word task. *Bilingualism: Language and Cognition*, *17*(3), 630–645. doi: 10.1017/S1366728913000539

Hernandez, A. E., Martinez, A., & Kohnert, K. (2000). In search of the language switch: An fmri study of picture naming in spanish–english bilinguals. *Brain and Language*, *73*(3), 421 - 431. Retrieved from `http://www.sciencedirect.com/science/article/pii/S0093934X99922782` doi: https://doi.org/10.1006/brln.1999.2278

Hollingshead, A. (1975). *Four factor index of social status.* New Haven: Yale University Department of Sociology.

Hopp, H., Vogelbacher, M., Kieseier, T., & Thoma, D. (2019). Bilingual advantages in early foreign language learning: Effects of the minority and the majority language. *Learning and Instruction*, *61*, 99 - 110. Retrieved from `http://www.sciencedirect.com/science/article/pii/S0959475217306199` doi: https://doi.org/10.1016/j.learninstruc.2019.02.001

Isel, F. (2005). Special issue in honour of elizabeth bates. first- and second-language processing: Cross-linguistic neurophysiological evidence. *Editions Modulaires Européennes, Langage et l'Homme*, 79-95.

Isel, F. (2017). *Atypies langagières de l'enfance à l'âge adulte : apports de la psycholinguistique et des neurosciences cognitives.*

Isel, F., Hahne, A., Maess, B., & D Friederici, A. (2007, 04). Neurodynamics of sentence interpretation: Erp evidence from french. *Biological psychology*, *74*, 337-46. doi: 10.1016/j.biopsycho.2006.09.003

Jarvis, S., & Pavlenko, A. (2008). *Crosslinguistic influence in language and cognition.* New York/London: Routledge.

J. Price, C. (1999, 12). A functional imaging study of translation and language switching. *Brain*, *122*, 2221-2235. doi: 10.1093/brain/122.12.2221

Jörgensen, N. (1976). *Meningsbyggnaden i talad svenska.* Lund: Studentlitteratur.

Kaan, E., & Swaab, T. (2003, 02). Repair, revision, and complexity in syntactic analysis: An electrophysiological differentiation. *Journal of cognitive neuroscience*, *15*, 98-110. doi: 10.1162/089892903321107855

Kovelman, I., Baker, S. A., & Petitto, L.-A. (2008). Bilingual and monolingual brains compared: A functional magnetic resonance imaging investigation of syntactic processing and a possible neural signature of bilingualism. *Journal of Cognitive Neuroscience*, *20*, 153-169.

Kovelman, I., H Shalinsky, M., S White, K., N Schmitt, S., S Berens, M., Paymer, N., & Petitto, L. A. (2008, 11). Dual language use in sign-speech bimodal bilinguals using fnirs brain-imaging. *Brain and language*, *109*, 112-23. doi: 10.1016/j.bandl.2008.09.008

Krause, F., & Lindemann, O. (2014). Expyriment: A Python library for cognitive and neuroscientific experiments. *Behavior Research Methods*, *46(2)*, 416-428. doi: 10.3758/s13428-013-0390-6

Kutas, M., & A. Hillyard, S. (1980, 02). Reading senseless sentences: brain potentials reflect semantic incongruity. science 207, 203-205. *Science (New York, N.Y.)*, *207*, 203-5. doi: 10.1126/science.7350657

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: finding meaning in the n400 component of the event-related brain potential (erp). *Annual review of psychology*, *62*, 621-47.

Makowski, D. (2016). Package 'neuropsychology': An r toolbox for psychologists, neuropsychologists and neuroscientists [Computer software manual]. Retrieved from https://github.com/neuropsychology/neuropsychology.R ((0.5.0))

MATLAB. (2017). *version 9.2.0.538062 (r2017a)*. Natick, Massachusetts: The MathWorks Inc.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*(1), 49 - 100. Retrieved from http://www.sciencedirect.com/science/article/pii/S001002859990734X doi: https://doi.org/10.1006/cogp.1999.0734

Moreno, S., Bialystok, E., Wodniecka, Z., & Alain, C. (2010). Conflict resolution in sentence processing by bilinguals. *Journal of Neurolinguistics*, *23*(6), 564 - 579. Retrieved from http://www.sciencedirect.com/science/article/pii/S0911604410000552 doi: https://doi.org/10.1016/j.jneuroling.2010.05.002

Navon, D. H. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, *9*, 353-383.

Nee, D. E., Wager, T. D., & Jonides, J. (2007, Mar 01). Interference resolution: Insights from a meta-analysis of neuroimaging tasks. *Cognitive, Affective, & Behavioral Neuroscience*, *7*(1), 1–17. Retrieved from https://doi.org/10.3758/CABN.7.1.1 doi: 10.3758/CABN.7.1.1

Neville, H., Nicol, J., Barss, A., Forster, K., & Garrett, M. (1991, 04). Syntactically based sentence processing classes: Evidence from event-related brain potentials. *Journal of cognitive neuroscience*, *3*, 151-65. doi: 10.1162/jocn.1991.3.2.151

Odlin, T. (2012, 11). Crosslinguistic influence in second language acquisition.. doi: 10.1002/9781405198431.wbeal0292

Oldfield, R. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*, 97-113.

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. (2011). FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. *Computational Intelligence and Neuroscience*, *2011 (2011)*. doi: 10.1155/2011/156869

Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, *31*(6), 785 - 806. Retrieved from http://www.sciencedirect.com/science/article/pii/0749596X9290039Z doi: https://doi.org/10.1016/0749-596X(92)90039-Z

Osterhout, L., & Mobley, L. A. (1995). Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language*, *34*(6), 739 - 773. Retrieved from http://www.sciencedirect.com/science/article/pii/S0749596X85710339 doi: https://doi.org/10.1006/jmla.1995.1033

Osterhout, L., & Nicol, J. (1999). On the distinctiveness, independence, and time course of the brain responses to syntactic and semantic anomalies. *Language and Cognitive Processes*, *14*(3), 283-317. Retrieved from https://doi.org/10.1080/016909699386310 doi: 10.1080/016909699386310

Paap, K. R., Anders-Jefferson, R., Mason, L., Alvarado, K., & Zimiga, B. (2018). Bilingual advantages in inhibition or selective attention: More challenges. *Frontiers in psychology*, *9*(1409). doi: 10.3389/fpsyg.2018.01409

Papageorgiou, A., Bright, P., Tomas, E. P., & Filippi, R. (2019). Evidence against a cognitive advantage in the older bilingual population. *Quarterly Journal of Experimental Psychology*, *72*(6), 1354-1363. Retrieved from `https://doi.org/10.1177/1747021818796475` (PMID: 30081734) doi: 10.1177/1747021818796475

Philipp, A., & Koch, I. (2009, 10). Inhibition in language switching: What is inhibited when switching between languages in naming tasks? *Journal of experimental psychology. Learning, memory, and cognition*, *35*, 1187-95. doi: 10.1037/a0016376

Philipp, A. M., Gade, M., & Koch, I. (2007). Inhibitory processes in language switching: Evidence from switching language-defined response sets. *European Journal of Cognitive Psychology*, *19*(3), 395-416. Retrieved from `https://doi.org/10.1080/09541440600758812` doi: 10.1080/09541440600758812

Psychology Software Tools, Inc. (2016). *E-Prime 2.0.* `http://www.pstnet.com`. Pittsburgh, PA.

R Core Team. (2019). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from `https://www.R-project.org/`

Rossum, G. (1995). *Python reference manual* (Tech. Rep.). Amsterdam, The Netherlands, The Netherlands.

Roussel, M., & Godefroy, O. (2008). La batterie grefex: donnees normatives. *Fonctions exécutives et pathologies neurologiques et psychiatriques*, 231-266.

Sharwood Smith, ., Michael, & Kellerman, E. (1986). *Crosslinguistic influence in second language acquisition* (1st ed ed.) [Book]. New York : Pergamon Institute of English. (Includes bibliographies)

Steinhauer, K., & Drury, J. E. (2012). On the early left-anterior negativity (elan) in syntax studies. *Brain and Language*, *120*(2), 135 - 162. Retrieved from `http://www.sciencedirect.com/science/article/pii/S0093934X11001246` (The Neurobiology of Syntax) doi: https://doi.org/10.1016/j.bandl.2011.07.001

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*, 643-662.

SurveyMonkey Inc. (2019). *Surveymonkey.* `www.surveymonkey.com`. San Mateo, California.

Swainson, R., Cunnington , R., & Jackson, S. (2001, 08). Erp correlates of executive control during repeated language switching. *Bilingualism: Language and Cognition*, *4*, 169 - 178. doi: 10.1017/S1366728901000268

Swedex. (2012). *Swedish examinations.*

Treccani, B., Argyri, E., Sorace, A., & Della Sala, S. (2009, Apr 01). Spatial negative priming in bilingualism. *Psychonomic Bulletin & Review*, *16*(2), 320–327. Retrieved from `https://doi.org/10.3758/PBR.16.2.320` doi: 10.3758/PBR.16.2.320

van Herten, M., Chwilla, D., & Kolk, H. (2006, 08). When heuristics clash with parsing routines: Erp evidence for conflict monitoring in sentence perception. *Journal of cognitive neuroscience*, *18*, 1181-97. doi: 10.1162/jocn.2006.18.7.1181

Wang, Y., Xue, G., Chen, C., Xue, F., & Dong, Q. (2007). Neural bases of asymmetric language switching in second-language learners: An er-fmri study. *NeuroImage*, *35*(2), 862 - 870. Retrieved from `http://www.sciencedirect.com/science/article/pii/S1053811906009694` doi: https://doi.org/10.1016/j.neuroimage.2006.09.054

Westman, M. (1974). *Bruksprosa.* Lund: Liber läromedel - Gleerup.

Weyerts, H., Penke, M., Münte, T. F., Heinze, H.-J., & Clahsen, H. (2002, May 01). Word order in sentence processing: An experimental study of verb placement in german. *Journal of Psycholinguistic Research*, *31*(3), 211–268. Retrieved from `https://doi.org/10.1023/A:1015588012457` doi: 10.1023/A:1015588012457

Wickens, T. D. (2001, 01). Elementary signal detection theory. *Elementary Signal Detection Theory*. doi: 10.1093/acprof:oso/9780195092509.001.0001

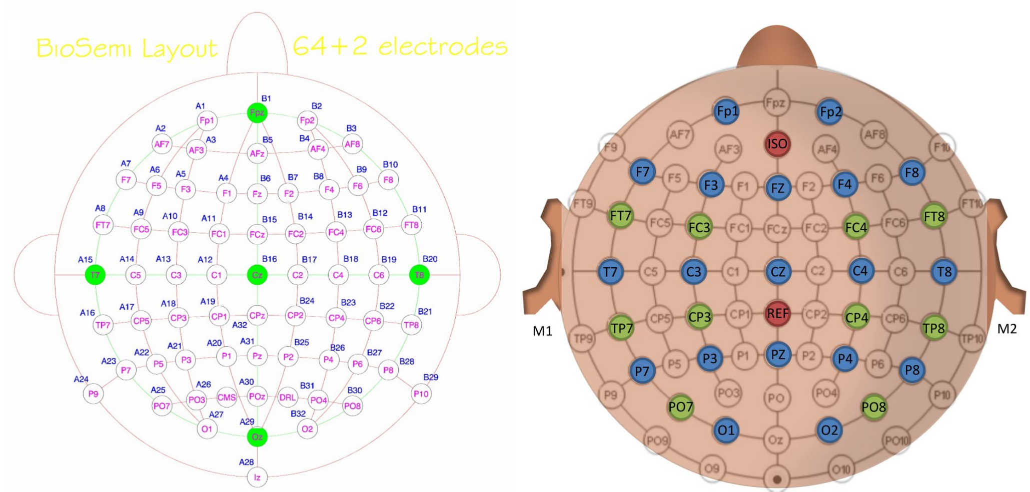
Wolleb, A., Sorace, A., & Westergaard, M. (2017, 01). Exploring the role of cognitive control in syntactic processing: Evidence from cross-language priming in bilingual children. *Linguistic Approaches to Bilingualism*, *8*. doi: 10.1075/lab.17002.wol

Ye, Z., & Zhou, X. (2009a). Conflict control during sentence comprehension: fmri evidence. *NeuroImage*, *48*(1), 280 - 290. Retrieved from http://www.sciencedirect.com/science/article/pii/S1053811909006582 doi: https://doi.org/10.1016/j.neuroimage.2009.06.032

Ye, Z., & Zhou, X. (2009b, 09). Executive control in language processing. *Neuroscience and biobehavioral reviews*, *33*, 1168-77. doi: 10.1016/j.neubiorev.2009.03.003

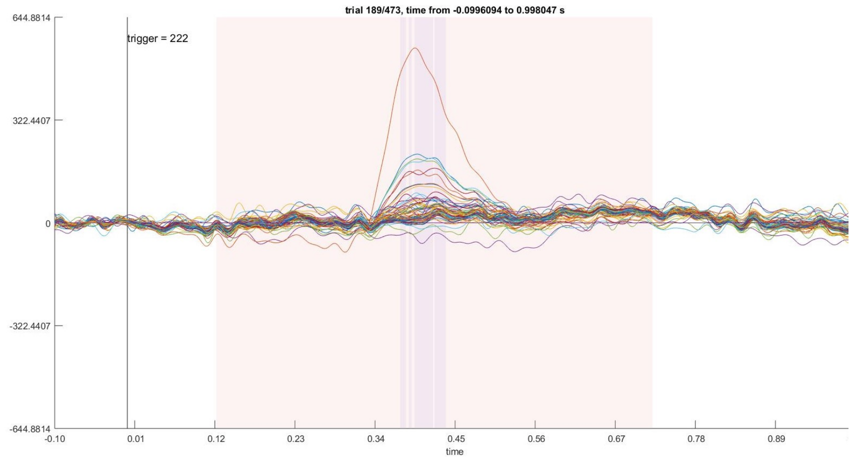
Zawiszewski, A., GUTIÉRREZ, E., FERNÁNDEZ, B., & Laka, I. (2011, 07). Language distance and non-native syntactic processing: Evidence from event-related potentials. *Bilingualism: Language and Cognition*, *14*, 400 - 411. doi: 10.1017/S1366728910000350

Appendix A Supplementary figures

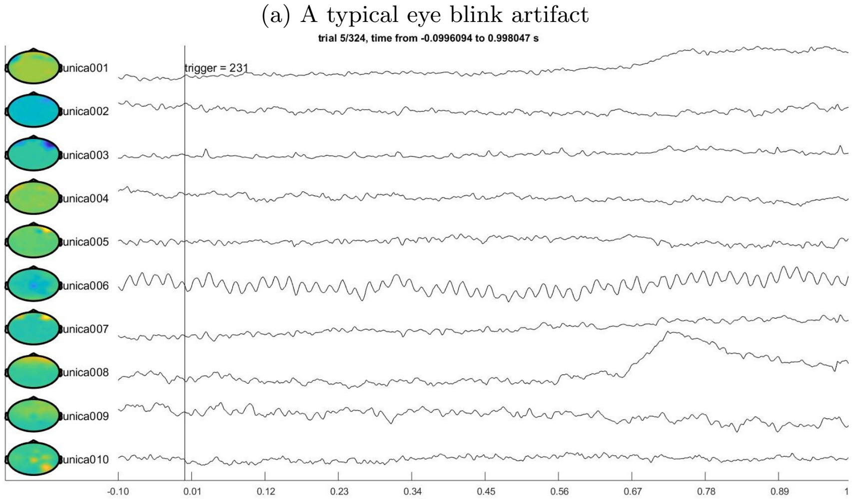


(a) All electrode sites in a 64-electrode array (b) Electrode sites used in ERP analyses

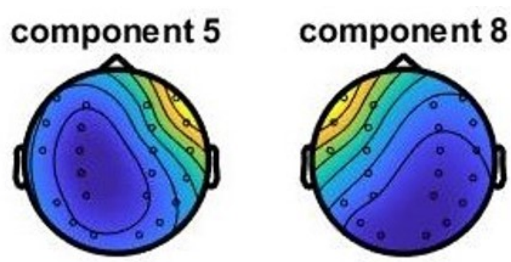
Figure 12: Electrode layouts under the International 10-20 System. Under this system, most electrodes are identified by a letter-number identifier. The letter represents the general part of the brain over which that electrode is placed, for example F for frontal, T for temporal, C for central, P for parietal, O for occipital. For electrodes falling at the margins of these regions, they get two-letter identifiers, e.g., FT for fronto-temporal. The number following the letter is assigned according to distance from the midline. The lower numbers are closer to the midline, while larger numbers are more lateral. Odd numbers are on the left side of the head, and even numbers on the right. Midline electrodes have the letter z instead of a number.



(a) A typical eye blink artifact



(b) Independent components after ICA decomposition. Note that components 1 and 8 reflect the mucular activity from an eye blink artifact.



(c) Typical topographies of eye blink components

Figure 13: Ocular artifacts. In order to work with clean data, it is necessary to remove all movement and ocular artifacts from the data while maintaining as much of the data as possible. There are two main methods to remove artifacts: totally discard trials, or use ICA to isolate and remove movement-based artifacts. In 13a, we see a typical eye blink artifact where the muscular activity for the blink is much greater than that of the neural activity around it. In 13b, we see the component time course for a trial that contained a blink. Note that the blink is captured mainly by component 8, but is also reflected in some of the other components. In 13c, we see typical topographies for ocular artifact-containing components. Note the pattern of activity localized on the front/side of the head while the rest of the topography does not seem to show any activity.

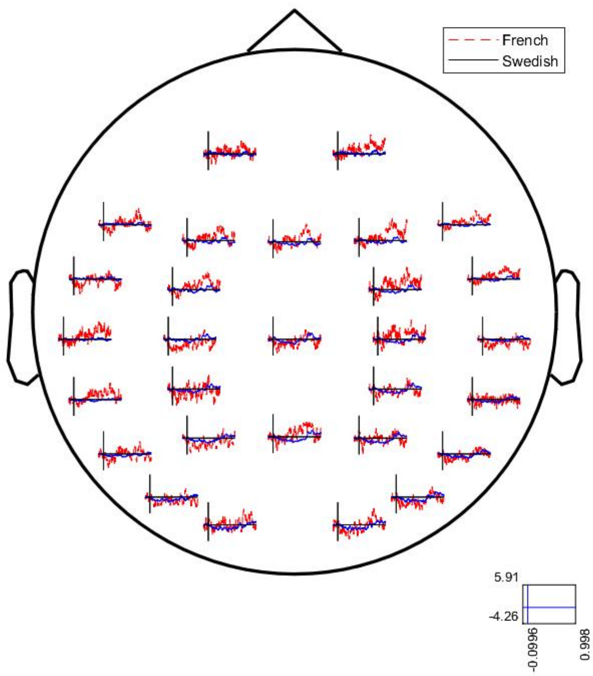


Figure 14: Difference ERPS for French and Swedish groups. We show here the difference (V3-V2) amplitude ERPs in the French and Swedish groups. In this plot, the two language groups are presented on the same scale, and the amplitude of the French ERP greatly overshadows the Swedish one. It is for this reason that they were not compared on the same axes throughout the rest of the paper. We believe this difference to be from differences of equipment, but plan to investigate further.

Appendix B Disclosures and consent form

<u>Notice d'informations</u>

Chers participants,

Nous sommes une équipe de chercheurs du laboratoire de linguistique MODYCO qui étudient les processus traitement du langage, et plus précisément chez les bilingues.
Nous faisons appel à vous ce jour pour une tâche de traitement de la syntaxe en suédois. L'expérience dure environ deux heures et demie. En plus de la tâche, il faudra compter 35 minutes d'installation et 5 minutes de désinstallation du système de mesure électroencéphalographique.
Des pauses seront comprises au sein de l'expérience.

Frédéric ISEL, Professeur des universités Université Paris Nanterre

Jeremy Yeaton, étudiant Master Sciences Cognitives, Ecole Normale Supérieure

Tâche de traitement de la syntaxe en Suédois

MODYCO UMR7114 CNRS & Université Paris Nanterre

<u>Autorisation pour l'enregistrement, l'archivage, l'exploitation scientifique et la diffusion de données</u>

Je soussigné(e)

Né(e) le :

Adresse :

Tel ou sms :
Mail :

- ☐ Accepte librement et en toute connaissance de cause de participer à cette recherche en électroencéphalographie (EEG) portant sur le traitement de syntaxe chez les bilingues.
- ☐ Il m'a été précisé que je suis libre d'accepter ou de refuser de participer à cette étude, et que mon nom n'apparaitra nulle part dans les publications.
- ☐ J'ai lu et pris connaissance des éléments contenus dans la lettre d'information. J'accepte de participer à cette étude dans les conditions qui y sont exposées. Mon consentement ne décharge en rien l'organisateur de cette étude de ses responsabilités et je conserve tous les droits qui me sont garantis par la loi.
- ☐ Je pourrai à tout moment mettre un ferme à ma participation à l'étude, sans en supporter aucune conséquence, ni avoir à me justifier.
- ☐ Les données recueillies resteront strictement confidentielles et l'anonymat est garanti.
- ☐ J'accepte que les données expérimentales recueillies à l'occasion de cette étude puissent faire l'objet d'un traitement informatise et soient communiques dans le cadre de publication dans des revues scientifiques.
- ☐ J'ai compris que toutes les données et résultats aux différentes épreuves seront anonymisés. Mes données personnelles (nom, date de naissance, adresse postale, adresse mail, téléphone) ne seront pas associées aux données scientifiques et ne seront pas diffusées.

Les membres du projet m'ont clairement expliqué les objectifs de ce projet, j'ai obtenu des réponses à mes questions. <u>À tout moment, j'ai le droit de demander que les enregistrements soient modifiés ou supprimés.</u>

Fait à le,

Signature, précédée de *lu et approuvé*

Pour toute information :

Jeremy YEATON
07 67 25 17 43
jdyeaton27@gmail.com