

Does better explicit knowledge of a morphosyntactic structure guarantee more native-like electrophysiological processing? An ERP study with French learners of English

Abstract. The possible transformation of the explicit knowledge developed during classroom learning of a second language into implicit knowledge of that language remains an open issue. In this study, we investigated the relationship between the explicit and implicit processing of morphosyntactic violations of English as an L2. ERP responses were obtained from 24 French learners (12 Intermediate and 12 Advanced) and 12 Native controls (NS) while participants evaluated the grammaticality of orally presented sentences containing subject-verb agreement violations. Results show that NS and Advanced speakers outperformed the Intermediate ones on the behavioural task. A P600 effect was obtained for all groups. Additionally, NS and Advanced learners exhibited an early negativity after violations while there was no significant effect in Intermediate speakers. The presence and amplitude of this early negativity was correlated with the structure-specific proficiency of Intermediate speakers and with the time of instruction of all learners. Results suggest that the superior native-likeness of the early responses obtained in Advanced learners is due more to their better proficiency and superior degree of explicit instruction than to the direct opportunity for implicit knowledge that their stay abroad represented.

1 Introduction

1.1 Implicit and Explicit Knowledge

Processing a language in real time relies on different types of knowledge that are often referred to as implicit and explicit knowledge (Andringa & Rebuschat, 2015; Rebuschat & Williams, 2013; Ullman, 2001). Implicit knowledge is usually characterised as knowledge we are unaware of and associated with automatic processing and procedural memory. It involves being able to complete a task without necessarily being aware of it and without having the capacity to explicitly describe how we do it. This is typically the case for walking or riding a bike, or, in terms of language abilities, for processing and producing the grammar of our first language. Explicit knowledge on the other hand is knowledge that we know we have, associated with certain types of conscious processes to retrieve specific memories. It is often linked with declarative memory and, when it comes to language, it is frequently associated with metalinguistic knowledge and the ability to describe the rules of a language. However, knowing the rules of a language does not mean one is fluent in it, and one can sometimes say more *about* the language they are learning than *in* it.

For most people, learning a second language in their home country implies learning it at school and therefore mostly through explicit instruction, usually with some focus on the rules of grammar. However, no consensus has been reached as to the degree to which the explicit and metalinguistic knowledge acquired in the classroom can be transformed into implicit knowledge and be integrated, automatized and proceduralised to be used without consciously recalling the explicit rules. Some researchers argue that no implicit knowledge can be transformed into explicit knowledge (no interface, e.g. Krashen, 1982), others claim that only some of it can and at the right developmental stage or that both types of knowledge can cooperate for learning (weak interface, e.g. R. Ellis, 1994 or N. Ellis, 2008), and others argue that all explicit knowledge can be transformed into implicit knowledge with time and practice (strong interface position, e.g. DeKeyser, 2007).

In this study, we were interested in the relationship between the explicit and the implicit processing of morphosyntactic violations in English as an L2 by French learners, and in the influence the proficiency of the participant can have on these two types of processing. To this purpose, we used two main measures: a grammaticality judgment task which requires some degree of conscious processing and is therefore associated with more explicit although not necessarily verbalised knowledge of syntactic rules, and event-related potentials (ERPs) which are automatic responses of the brain and thus commonly linked to unconscious knowledge and processing.

We focused on a structure that works in a similar way in French and English and is thus likely to benefit from positive transfer: subject-verb agreement. Subject-verb agreement in French is realized orally through different inflectional suffixes (from 3 to 5 different endings depending on the type of verb); in English, its only instance is in the 3rd person singular in the indicative present tense. This subject-agreement rule is taught very early and repeatedly throughout the years of English instruction French students complete. Yet omitting the 3rd person *-s* is one of the most common mistakes learners make and one of the most persistent ones, which suggests that this grammatical rule is resistant to instruction and that having good explicit knowledge of this rule does not directly influence the way it is processed. Despite the positive transfer and the apparent simplicity of the rule, proceduralisation of this rule seems to be problematic and suggest a need for more implicit knowledge.

1.2 ERPs

ERPs represent changes in brain electric activity triggered by particular events (Fabiani, Gratton, & Federmeier, 2007; van Hell & Tokowicz, 2010). Two main ERP components have been identified for the study of syntactic processing (Hahne & Friederici, 1999, 2001) and are of interest for this study. The first one is the Left Anterior Negativity (LAN), a negative shift occurring between 300 and 500 ms after morphosyntactic violations (Chen, Shu, Liu, Zhao, & Li, 2007). It is said to evidence automatic processing of morphosyntax (Friederici, 2002). It is also sometimes preceded by an Early Left Anterior Negativity (ELAN) (Friederici, 2002; Hahne & Friederici, 1999). The P600, a positive shift maximal at centro-parietal locations (van Hell & Tokowicz, 2010) between 500 and 900 ms, has been consistently observed in response to a large variety of syntactic violations (Bond, Gabriele, Fiorentino & Alemán Bañon, 2011; Hahne & Friederici, 2001; Meulman, Stowe, Sprenger, Bresser, & Schmid, 2014; Tanner, McLaughlin, Herschensohn, & Osterhout, 2013). It is thought to reflect control and reanalysis processes (Kaan, Harris, Gibson, & Holcomb, 2000).

ERPs observed in native and non-native speakers often differ, either in a qualitative (different component, Tanner, McLaughlin, et al., 2013) or a quantitative way (delays and amplitude reductions, Sabourin & Stowe, 2008). Several factors have been identified to explain why some learners exhibit similar ERPs as native speakers and others do not, among

which proficiency – which has been shown to positively correlate with the magnitude of the observed ERP (Ojima, Nakata, & Kakigi, 2005; Rossi, Gugler, Friederici, & Hahne, 2006; Tanner, Inoue, & Osterhout, 2013). Proficiency was found to induce quantitative (delayed P600s with reduced amplitude in intermediate compared to advanced learners in Tanner, Inoue, et al., 2013; Tanner, McLaughlin, et al., 2013) and qualitative differences (a LAN in advanced speakers only in Rossi et al., 2006). Since our French participants have different proficiency levels, we expect to find these differences between our groups and/or with the native speakers.

ERP components also differ in terms of their relation with implicitness. Although early negativities like the ELAN and the LAN are deemed automatic and thus require a certain degree of implicit processing, later components like the P600 reflect controlled processes and are not found in the absence of attention (Pulvermüller, Shtyrov, Hasting, & Carlyon, 2008), which suggests that they require explicit processing.

1.3 Research hypotheses

French learners who participated in this study were divided into 2 groups based on their year of University instruction. Due to the omnipresence of the subject-verb agreement rule in English instruction, it is expected that learners in both groups had similar explicit knowledge of this particular structure. No major differences in later ERP components (P600) were expected between the groups. However, participants who were more advanced were expected to have better explicit knowledge of the language in general. Since processing of the auditory stimuli involved paying attention to more than just the subject-verb agreement rule, differences in the behavioural performance (measured by the sensitivity index or d') could still be observed.

Participants in the Advanced group had spent an academic year abroad. This should be associated with the development of more implicit knowledge, especially for a structure that is so common. However the impact of study abroad experiences on the development of sensitivity to grammatical violations is usually very limited or null (Segalowitz, 2004). If a 1-year study abroad experience is enough for the development of implicit knowledge, we should see differences in early ERPs (ELAN or LAN) between the Advanced and Intermediate learners.

For Native Speakers, we expect to find ERP results similar to those obtained in previous studies: a possible LAN followed by a P600. Their behavioural performance should be superior to that of Intermediate learners, but could be statistically indistinguishable from that of Advanced learners.

2 Methods

2.1 Participants

24 French learners of English as well as 12 Native Speakers took part in the experiment. All were right-handed (as assessed by the Edinburgh Handedness Inventory (Oldfield, 1971)) with normal or corrected sight and normal hearing (as confirmed by a hearing test). All participants were paid for their participation. Native Speakers of English had grown up in England or the United States of America and were exchange students in a French University. French participants were English Majors at one of the Universities in Lyon. They were divided into two groups based on their degree and a proficiency test completed prior to the experiment targeting oral comprehension and production. 12 of them were first-year students who had an intermediate level of English and had never spent more than 2 weeks in an

English-speaking country. The remaining 12 were graduates who had spent one academic year in an English-speaking country and had an advanced level of proficiency. They were Master’s students and many of them were preparing for teaching certifications (CAPES). Mean ages for each group are reported in Table 1.

Table 1. Number and Mean Age of Participants per Group.

| Group | Number of participants (male participants) | Mean age | Standard Deviation |
|-----------------------|---|----------|-----------------------|
| Native Speakers | 12 (1) | 21y10m | 1y10m |
| Advanced Learners | 12 (3) | 23y10m | 1y2m |
| Intermediate Learners | 12 (5) | 20y | 1y2m |

The two groups of learners differed on a number of points, the first one being their proficiency, evaluated according to the CEFR (Common European Framework of Reference for Languages): participants in the Intermediate group had a B1 to B2 level (“independent user”) whereas participants in the Advanced group had a C1 to C2 level (“proficient user”). Their degree of exposure to the language (2 weeks at most vs. 1 year) and their time of University instruction (1 semester vs. 5 to 7 semesters) differed as well. There was a significant difference between their total time of instruction ($t(21.99)=3.77$, $p<.01$; $\text{Mean}_{\text{Advanced}} = 13\text{y}8\text{m} \pm 22\text{m}$; $\text{Mean}_{\text{Intermediate}} = 10\text{y}9\text{m} \pm 23\text{m}$). Finally, the two groups differed significantly in their self-assessment of their overall English proficiency ($t(16.25)=2.22$, $p<.05$; $\text{Mean}_{\text{Advanced}} = 16.75/20 \pm 2.09$; $\text{Mean}_{\text{Intermediate}} = 15.25/20 \pm 1.06$). More precisely, the differences concerned their assessment of their own abilities in oral comprehension ($\text{Mean}_{\text{Advanced}} = 4.25/5 \pm 0.45$; $\text{Mean}_{\text{Intermediate}} = 3.75/5 \pm 1.06$; $t(20.10)=2.25$, $p<.05$), which was a key parameter for this experiment since stimuli were presented orally.

2.2 Material

The critical material consisted of 160 short active sentences composed of a 3rd person pronoun (*He* or *They*), a verb in the present tense and a short complement (adjective and noun or frequent adverbial phrase). All the lexical words used in these sentences were among the 1000 most frequent words of their grammatical category in the COCA (*Corpus of Contemporary American English*, the largest corpus of contemporary English that is freely available and the only balanced corpus of American English, see Davies, 2008), so as to reduce possible difficulties in understanding the sentences for the participants in the Intermediate group. Words specific to American English were avoided so as not to unsettle the participants who were native speakers of British English.

Half of the sentences contained a violation of Subject-Verb agreement – the presence or absence of the *-s* at the end of the main verb determined the grammatical acceptability of the sentence. Half of the sentences were constructed with the pronoun *He* and half with the pronoun *They*, which enabled us to balance the presence of the *-s* in correct and incorrect sentences (see examples (1) and (2)).

- (1) a. *They reject the proposition.*
PRO.3PL rejeta.PST la proposition
‘Ils rejettent la proposition.’
- b. **They rejects the proposition.*
PRO.3PL rejeta.PST.3SG la proposition

‘Ils rejettent la proposition.’

- (2) a. *He meets people every day.*
PRO.3SG rencontrer.PST.3SG gens tous jours.
‘Il rencontre des gens tous les jours.’
- b. **He meet people every day.*
PRO.3SG rencontrer.PST gens tous jours.
‘Il rencontre des gens tous les jours.’

All the stimuli were recorded by a native speaker of English (from Portland, Oregon, USA) at a natural pace with an Audio Technica AT2020 USB microphone in a sound-attenuated booth with ROCme! software (Ferragne, Flavier, & Fressard, 2012). Stimuli were digitized as PCM mono, 44.1 kHz, 16 bit. Irregularities in the frequency response of the microphone and headphones (BeyerDynamics DT770 Pro – 250 Ohms) were corrected with the equalization function in Audacity (Team Audacity, 2013). RMS intensity was made equal across stimuli with the Praat program (Boersma & Weenink, 2007). The critical point – the instant from which event-related potentials were recorded – was the beginning of the word following the violation, meaning the moment at which the sentence becomes correct or incorrect.

2.3 Procedure

2.3.1 Preliminary tasks

Participants first completed a language background questionnaire collecting data on the different languages they spoke and their motivation to learn foreign languages. The questionnaire for French participants, inspired from standard questionnaires used in second language acquisition (Gullberg & Indefrey, 2003; Marian, Blumenfeld, & Kaushanskaya, 2007; Ping Li, 2006), also contained questions about their age of acquisition of English, the time they had spent in an English-speaking country and a self-assessment of their proficiency.

Participants’ hearing was tested with an Electronica Technologies 600M USB audiometer. The Hughson-Westlake method (Carhart & Jerger, 1959) was used to determine their hearing threshold, which was confirmed to be normal.

French participants also completed a proficiency test including a listening and a speaking task. The oral comprehension task was based on the test these participants had completed for their Baccalauréat. It enabled us to make sure that the participants were proficient enough to understand the audio stimuli in the main experiment. Participants listened to an audio document (a 1-minute extract from a BBC podcast) twice and gave an account of it in French. The speaking test was inspired by TOEFL tasks. Participants had 15 seconds to prepare for their production and then were asked to speak for 45 to 60 seconds about a given topic. They completed this task twice: first to tell an amusing anecdote from their personal life and then to express and justify an opinion on a specific topic.

2.3.2 Main experiment

For the main part of the experiment, participants listened to the recorded stimuli in a sound-attenuated room with dim lighting. Stimuli were presented with Presentation software (Neurobehavioral Systems, 2012) at a mean intensity of 70 dBA (calibrated with a GRAS 43 AG artificial ear) with the following procedure (see Figure 1). A fixation cross appeared first

for 1000 ms and remained on the screen during the auditory presentation of the stimulus (around 1500 to 2000 ms) to reduce ocular movements and for 1000 ms after the end of the stimulus. A screen prompted the participant to evaluate the grammaticality of the sentence and remained for at most 2000 ms. Participants answered by pressing a coloured button on a response box.

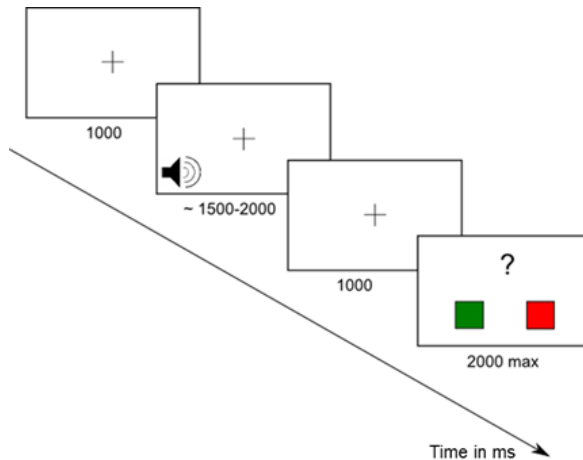


Figure 1. Experimental Procedure

2.3.3 EEG data acquisition

EEGs were recorded with a Biosemi ActiveTwo system with 32 active electrodes at the following sites: Fp1, Fp2, F3, Fz, F4, F8, FC5, FC1, FC2, FC6, T7, C3, Cz, C4, T8, CP5, CP1, CP2, CP6, P7, P3, Pz, P4, P8, PO3, PO4, O1, Oz and O2. Vertical and horizontal electro-oculograms were recorded with a bipolar montage using electrodes placed above, underneath and to the left of the left eye. EEGs were referenced on-line to the average of all electrodes and re-referenced off-line to the average of the two mastoids. Data were filtered on-line between 0.1 and 100 Hz. Electrode offset was maintained below 20 mV and the signal was sampled at a rate of 512 Hz.

2.3.4 Data analysis

Hit rates (correct acceptance of grammatical sentences) and false alarms (incorrect acceptance of ungrammatical sentences) were computed for each subject according to signal detection theory (J. Tanner, Wilson, & Swets, 1954). The response variable used in analyses of behavioural data was the sensitivity index (d') of the participants. A linear mixed-effect model with d' as a dependent variable and Group (Native/Intermediate/Advanced) and Pronoun (*He/They*) as categorical predictors was conducted with the R program (R Core Team, 2014).

Since the stimuli were followed by a 1000 ms interval, response times were not precise enough to reveal processing difficulties and were therefore not analysed. All trials were

included in the ERPs analysis, independently of the accuracy of the behavioural response¹. EEG data were analysed with MATLAB (The Mathworks, 2008) and the EEGLAB (Delorme & Makeig, 2004) and ERPLAB (Lopez-Calderón & Luck, 2014) toolboxes. Epochs from -200 ms to 900 ms around the critical point were extracted from continuous data. After baseline correction (-200-0 ms), high-pass filtering at 0.1 Hz and low-pass filtering at 30 Hz, trials for which peak-to-peak amplitude exceeded 70 μV on the EOG channel or 150 μV on the other channels were automatically rejected. Visual inspection completed the artefact rejection process. ERPs were averaged for each subject and each group. Electrodes selected for the analyses were divided into central line electrodes (Fz, Cz, Pz) and lateral electrodes, themselves divided into anterior/central/posterior region and left/right hemisphere (anterior left: F3, F7, FP1, AF3; anterior right: F4, F8, FP2, AF4; central left: FC1, FC5, C3, CP1, CP5; central right: FC2, FC6, C4, CP2, CP6; posterior left: P3, P7, PO3, O1; posterior right: P4, P8, PO4, O2).

According to the literature and after visual inspection of ERP waveforms, several temporal windows of interest were selected: a P600 window (500-900 ms) and two windows for early negativities: 100-300 ms and 300-500 ms. The mean amplitude in each window was computed for each subject and each condition and for the difference between Incorrect and Correct conditions.

Results were analysed using linear mixed-effect models in R version 3.1.0, packages lme4 version 1.1.10 (Bates et al., 2015), LMERConvenienceFunctions version 2.10 (Tremblay & Ransijn, 2015), lmerTest version 2.0-33 (Kuznetsova, Brockhoff, & Christensen, 2015), and lsmeans version 2.25 (Lenth, 2016). Each component was analysed in five steps:

1. An initial model was built with the maximal fixed-effect structure and only a random intercept. The dependent variable was the mean amplitude in the selected time-window.
2. The best random-effect structure was determined by forward-fitting the random effect structure through log-likelihood ratio comparisons.
(LMERConvenienceFunctions::ffRanefLMER.fnc)
3. The resulting model was reduced by removing non-significant higher order fixed effects through log-likelihood comparisons
(LMERConvenienceFunctions::bgFixefLMER_F.fnc)
4. The resulting model was submitted to a type III Anova with a Satterthwaite approximation of the degrees of freedom (lmerTest::anova)
5. Higher order interactions were followed-up with pairwise post-hoc tests adjusted for multiple comparisons with Tukey's honest significant difference (lsmeans::lsmeans)

Since the factor Region was always included in the highest order interaction, we followed-up these interactions by building different models for each level of Region: Anterior, Central and Posterior. This enabled us to directly compare the effects of Pronoun and Group in the same model. For brevity, only these models and only significant results are reported here.

The initial model thus included as initial fixed effects: Condition (Cond): Correct/Incorrect, Pronoun (Pro): *He/They*, Hemisphere (Hem): Left/Right, and Group (Group): Advanced/Intermediate/Native Speakers, and their interactions. The random effects evaluated for inclusion in the model were: Condition, Pronoun, Region and Hemisphere.

3 Results

¹ This is a recurring approach in second language acquisition (Chen et al., 2007; Mueller, 2005; Ojima et al., 2005; Weber-Fox & Neville, 1996) ; besides, some studies have found ERP effects without behavioural effects (Tokowicz & MacWhinney, 2005).

3.1 Behavioural data

Overall, participants detected violations successfully (see Table 2 for the mean d' and accuracy per group). Linear mixed-effect model analyses revealed a main effect of Group ($F(2,33) = 6.07, p < .01$), see also Figure 2. Pairwise comparisons (adjusted with Tukey's Honest Significant Difference) showed that accuracy was significantly higher for the Native Speakers (NS) than for the Intermediate Learners (IL) ($t(33)=3.39, p<.01$) and marginally higher for the Advanced Learners (AL) than the Intermediate ones ($t(33)=2.40, p=.06$). There was however no significant difference between NS and AL. There was also a main effect of Pronoun ($F(1,33)=5.47, p<.05$) due to the fact that participants all performed slightly better for violations with the pronoun *He* ($t(33)=2.34, p<.05$): their d' was on average 0.32 points better with *He* than with *They*.

Table 2. Mean d' and accuracy per Group on the grammaticality judgment task

| | d' | Accuracy |
|------------------------------|------|----------|
| Native Speakers | 4.13 | 97% |
| Advanced Learners | 3.78 | 94% |
| Intermediate Learners | 2.92 | 88% |

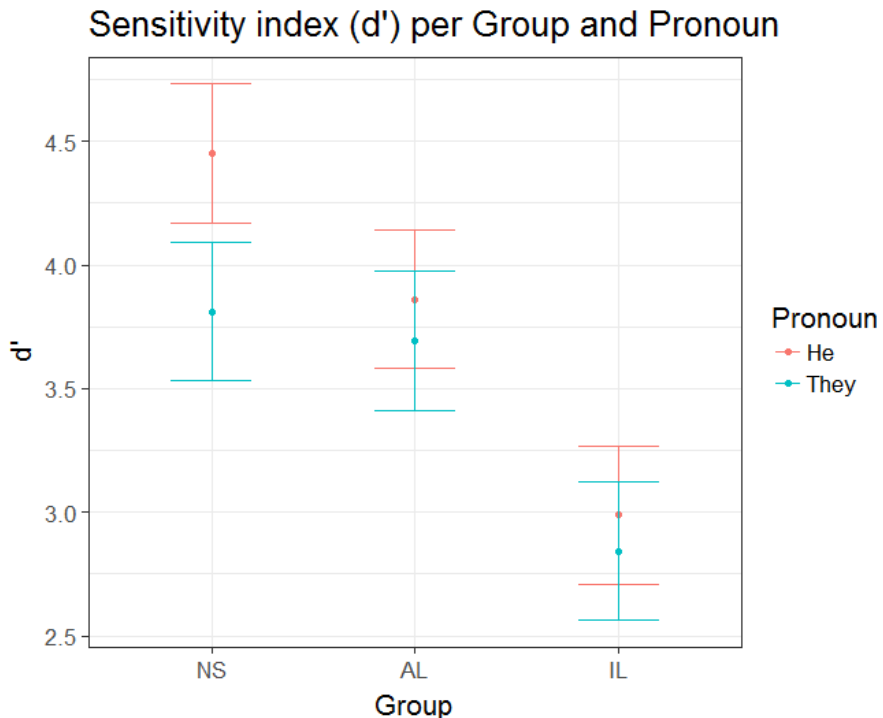


Figure 2. Sensitivity index (d') per Group and Pronoun (*Mean and Standard Deviation*).

3.2 EEG data

Waveforms for the Incorrect and Correct condition per Group and Pronoun can be seen in Figure 3.

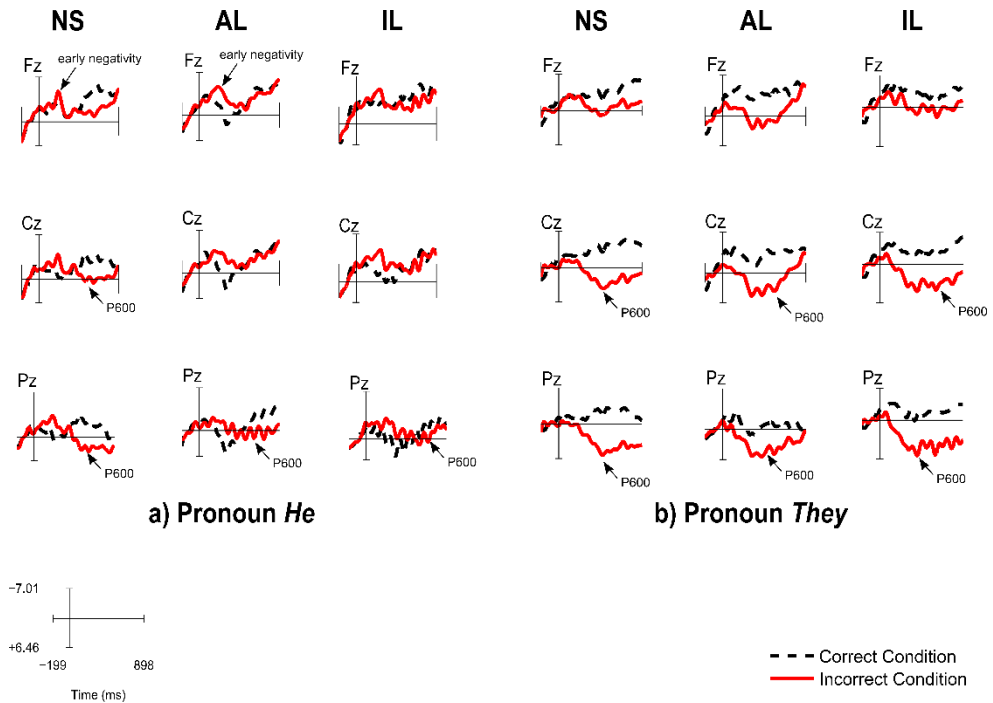


Figure 3. ERPs per Group and Pronoun

3.2.1 P600 Effect (500-900 ms)

Anterior. The Cond \times Pro interaction only was significant ($F(1,1147)=17.69, p=.000$). Post-hoc comparisons show that there was no effect of condition with the pronoun *He*, but a positivity with the pronoun *They* ($\text{Mean}_{(\text{TheyI-TheyC})}=1.57 \mu\text{V}, t(45.78)=4.1, p<.001$).

Central. Only the Cond \times Pro interaction was significant ($F(1,1295)=67.77, p=.000$). A positivity was found for both pronouns ($\text{Mean}_{(\text{HeI-HeC})}=1.08 \mu\text{V}, t(45.21)=2.9, p<.01$; $\text{Mean}_{(\text{TheyI-TheyC})}=3.19 \mu\text{V}, t(45.21)=8.6, p<.001$) but was larger with the pronoun *They* ($\text{Mean}_{((\text{TheyI-TheyC})-(\text{HeI-HeC}))}=2.06 \mu\text{V}, t(35)=3.6, p=.001$).

Posterior. The Cond \times Pro interaction only was significant ($F(1,1148)=30.87, p=.000$). The amplitude in the Incorrect condition was larger than in the Correct condition for both pronouns ($\text{Mean}_{(\text{HeI-HeC})}=1.16, t(47.35)=3.0 \mu\text{V}, p<.01$; $\text{Mean}_{(\text{TheyI-TheyC})}=2.76 \mu\text{V}, t(47.35)=7.2, p<.0001$). The difference between the two conditions was once again larger

with *They* ($\text{Mean}_{((\text{TheyI}-\text{TheyC})-(\text{HeI}-\text{HeC}))}=1.60 \mu\text{V}$, $t(35)=3.4$, $p<.01$) although less so than in the Central region.

In the 500-900 ms time window, violations elicited a broadly-distributed positivity with *They* and a Centro-Posterior one with *He* (thus resembling more a typical P600-effect) for all groups (see Figure 4).

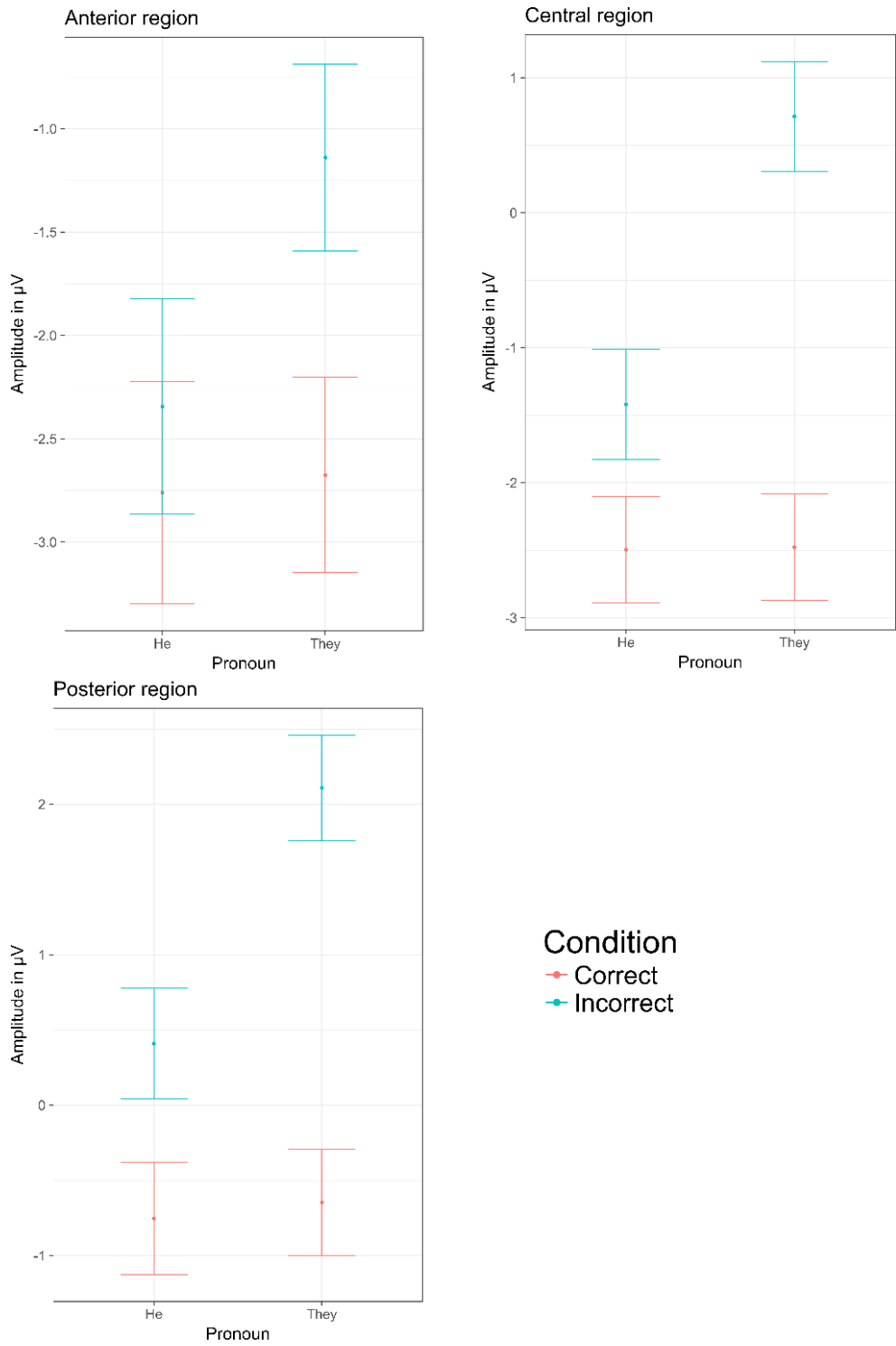


Figure 4. Amplitude in the 500-900 ms window per Condition, Pronoun and Region (*Mean and Standard Deviation*)

3.2.2 Early negativity (100-300 ms)

Anterior. The Cond \times Pro \times Group interaction almost reached significance ($F(2,1073)=2.99$, $p=.050$). Due to the level of the p value, this interaction was followed up by post-hoc comparisons. A negativity was found with the pronoun He for participants in the Advanced Group ($\text{Mean}_{(\text{HeI-HeC})}=-1.38$ μV , $t(1073)=5.2$, $p<.0001$) and the Native Speakers Group ($\text{Mean}_{(\text{HeI-HeC})}=-0.97$, $t(1073)=3.7$, $p<.01$) but not the Intermediate Group ($\text{Mean}_{(\text{HeI-HeC})}=-0.24$ μV , $t(1073)=0.9$, $p=.94$). Violations with the pronoun They elicited a positivity for participants in the Advanced ($\text{Mean}_{(\text{TheyI-TheyC})}=1.27$ μV , $t(1073)=4.8$, $p<.0001$) and Intermediate ($\text{Mean}_{(\text{TheyI-TheyC})}=1.44$ μV , $t(1073)=5.5$, $p<.0001$) Groups but not for Native Speakers ($\text{Mean}_{(\text{TheyI-TheyC})}=0.45$ μV , $t(1073)=1.7$, $p=.52$). Post-hoc comparisons for the two next significant higher order interactions Cond \times Pro ($F(1,1073)=79.14$, $p=.0000$) and Cond \times Group ($F(2,1073)=5.80$, $p=.003$) revealed an overall negativity with the pronoun He ($\text{Mean}_{(\text{HeI-HeC})}=-0.86$ μV , $t(1073)=5.67$, $p<.0001$) and a positivity with the pronoun They ($\text{Mean}_{(\text{TheyI-TheyC})}=1.05$ μV , $t(1073)=6.91$, $p<.0001$); as well as a general positivity for participants in the Intermediate Group ($\text{Mean}_{(\text{I-C})}=0.60$, $t(1073)=3.21$, $p<.01$) but nothing for the two other groups. This is not surprising since these results are averaged over pronouns and we know that for participants in the Advanced and Native Speakers groups, the polarity of the difference between Incorrect and Correct conditions was opposed between the two pronouns.

Central. Post-hoc comparisons for the significant Cond \times Pro \times Group interaction ($F(2,1427)=8.51$, $p=.0002$) revealed a negativity with He for participants in the Advanced Group only ($\text{Mean}_{(\text{HeI-HeC})}=-1.86$ μV , $t(42.86)=-4.2$, $p<.01$). There was a positivity with They for both learner groups ($\text{Mean}_{(\text{ALTheyI-ALTheyC})}=2.18$, $t(42.86)=4.9$, $p<.001$; $\text{Mean}_{(\text{ILTheyI-ILTheyC})}=2.46$, $t(42.86)=5.55$, $p<.0001$).

Posterior. There was a significant Cond \times Pro interaction ($F(1,1148)=166.46$, $p=.0000$). Violations with He triggered a negativity ($\text{Mean}_{(\text{HeI-HeC})}=-0.94$, $t(44.83)=-3.4$, $p<.01$) and violations with They elicited a positivity ($\text{Mean}_{(\text{TheyI-TheyC})}=1.51$, $t(44.83)=5.4$, $p<.0001$). Violations with the pronoun They thus triggered an early broadly-distributed positivity for both learner groups and a posterior one for Native Speakers. Violations with He were followed by an early anterior negativity for Native Speakers, a posterior negativity for Intermediate learners and a largely distributed one for Advanced Speakers (see Figure 5).

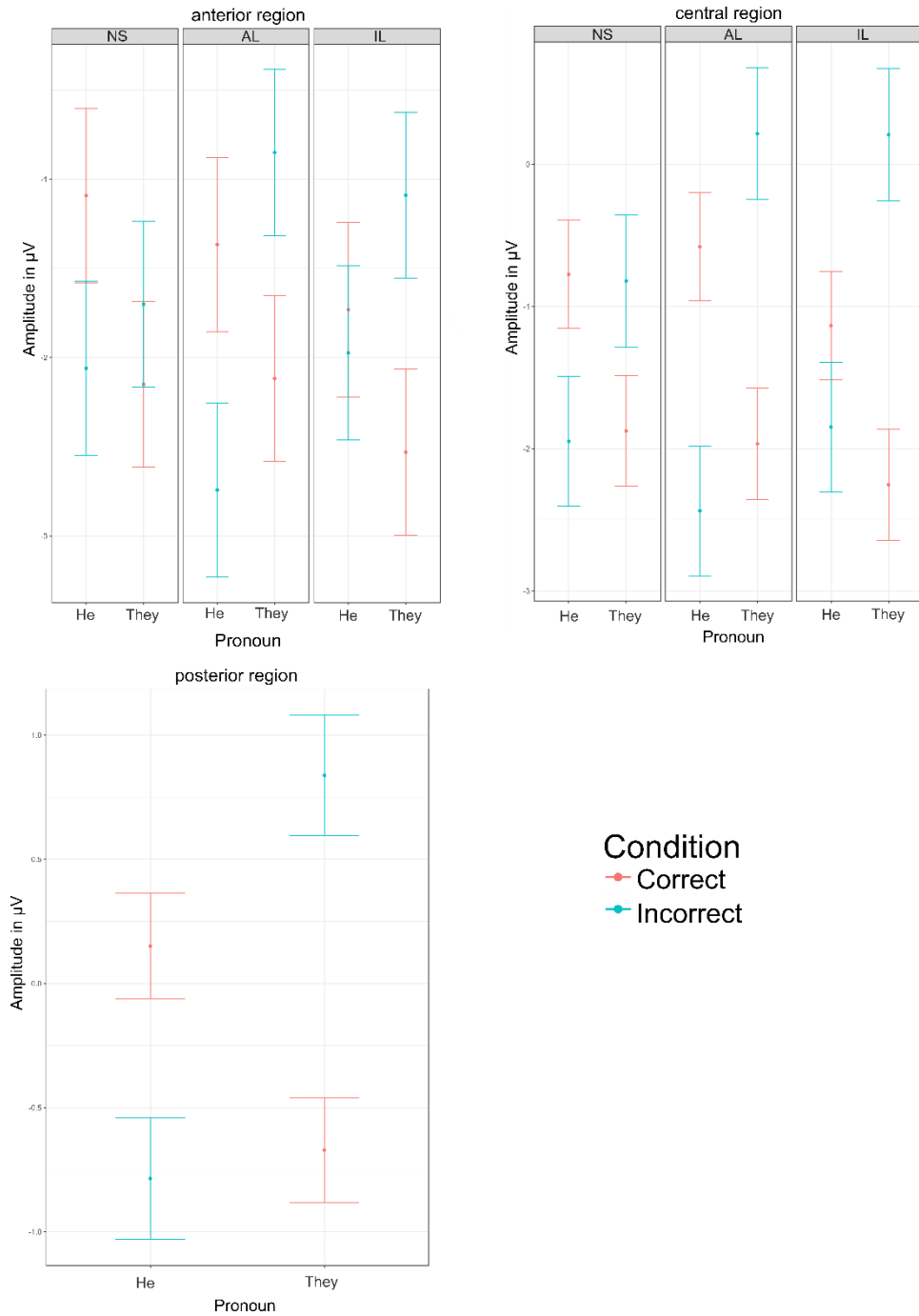


Figure 5. Amplitude in the 100-300 ms window per Condition, Group, Pronoun and Region (*Mean and Standard Deviation*)

3.2.3 Intermediate negativity / N400 (300-500 ms)

Anterior. The Cond \times Pro \times Gp was significant ($F(2,1139)=3.39, p=.03$). No effect was found with *He*; but with *They* the amplitude in the Incorrect condition was larger than in the Correct conditions for all groups of participants ($\text{Mean}_{(\text{ALTheyI-ALTheyC})}=1.70 \mu\text{V}, t(44.45)=3.0, p=.048$; $\text{Mean}_{(\text{ILTheyI-ILTheyC})}=2.17 \mu\text{V}, t(44.45)=3.8, p=.005$; $\text{Mean}_{(\text{NSTheyI-NSTheyC})}=2.02 \mu\text{V}, t(44.45)=3.56, p=.011$).

Central. The Cond \times Pro interaction was significant ($F(2,1295)=216.16, p=.0000$). There was no effect with *He*, but a positive one with *They* ($\text{Mean}_{(\text{TheyI-TheyC})}=3.61, t(52.01)=11.3, p<.0001$).

Posterior. The Cond \times Pron interaction was also significant ($F(2,1148)=170.27, p=.0000$). Post-hoc comparisons showed the same effects as in the central region: no effect with *He* and a general positivity with *They* ($\text{Mean}_{(\text{TheyI-TheyC})}=3.02, t(47.89)=9.0, p<.0001$).

In this window, violations with *They* triggered a broadly-distributed positivity for participants in all groups, while no effect followed violations with *He*.

3.3 Relationship between measures of explicit and implicit processing

3.3.1 Time of instruction and time spent abroad

Given the difference observed between the two pronouns, data was examined separately for each; and separate models were built for each region. The link between each component and the Time of instruction (which was not exactly identical to the Group separation); and between each component and the Time spent in an English-speaking country (which varied only for the Advanced Group) were examined.

There was a link between the mean amplitude between 100 and 300 ms and the Time of instruction in the anterior ($F(1,20)=4.58, p=.044$) and central regions ($F(1,20)=4.39, p=.049$) but not in the posterior one: the longer the participants had been learning English, the more negative-going the effect was.

There was no significant link between the amplitude of any component and the Time spent in an English-speaking country.

3.3.2 Proficiency

Models were built for each effect (Incorrect – Correct difference in each time window of interest) and each region with the d' as a continuous predictor and the Group as a categorical one. A random intercept was added in the model as well as a random slope for each of the repeated measures. Only significant results are reported here.

P600 effect. The $d' \times$ Group interaction was significant in the central region ($F(2,49.92)=3.48, p=.039$). This interaction was followed-up by separate group models. This

revealed that the effect of d' was only significant for Advanced learners ($F(1,13.54)=12.45, p=.004$): the larger the d' , the larger the amplitude of the P600 effect.

Early negativities. Because of the opposite effects the two pronouns had for several groups of learners, data was examined both generally and separately for each pronoun.

In the anterior region, the $d' \times$ Group interaction was marginally significant ($F(2,31.34)=2.50, p=.098$). This interaction was however significant for the pronoun *He* ($F(2,32.61)=5.28, p=.01$). The effect of d' was only significant for the Intermediate Group ($F(2,28.09)=9.62, p=.0007$): the larger the d' , the more negative-going the effect.

The $d' \times$ Group interaction was also marginally significant in the central region ($F(2,32.37)=3.18, p=.055$) and significant with the pronoun *He* ($F(2,29.54)=3.61, p=.04$). Follow-ups with *He* were not significant. Follow-ups of the general model found a significant effect of d' for Intermediate learners ($F(1,11.70)=6.02, p=.031$) and Native Speakers ($F(1,23.15)=4.50, p=.045$). These results have to be taken with caution but show that for these participants, there was a trend for the amplitude of the effect to decrease from a positivity towards a null difference as the d' increased.

300-500 ms window. The $d' \times$ Group interaction was marginally significant in the anterior region only ($F(2,47.21)=2.47, p=.095$). Intermediate learners tended to have a reduced positivity in this region when their d' increased ($F(1,10.54)=7.77, p=.02$).

These results show that the positivity that was found in earlier time windows tended to decrease with better proficiency or even to transform into the native-like negativity in the very early window.

4 Discussion

In this study, ERP responses to subject-verb agreement violations were obtained from two groups of French learners of English with different proficiency levels and a Native Speakers control group while participants evaluated the grammaticality of orally presented stimuli. Results show that Natives and Advanced learners outperformed the Intermediate learners on the behavioural task. Violations elicited a P600 for participants in all groups, but this P600 was more narrowly localized to the centro-posterior region (and thus more typical) with the pronoun *He*. The size of the effect depended on the structure-specific proficiency for Advanced learners, which is consistent with what has been found in other studies (e.g. Rossi et al., 2006). It is surprising however that this link was not also found for learners in the lower-proficiency group.

Both Native Speakers and Advanced learners exhibited an early anterior negativity suggesting actual automatic and implicit morphosyntactic processing of the violations with *He*; but violations with *They* triggered a broadly distributed positivity in both learner groups and a posterior positivity in Native speakers resembling a P3, an attention-related component. Advanced learners' most automatic responses were thus more native-like than those of Intermediate speakers. The positivity observed with the pronoun *They* could be due to the fact that these violations are phonologically more salient than the violations with *He* since they involve adding the $-s$ morpheme instead of removing it. They are also much less common. Indeed, the omission of the 3rd person singular $-s$ in the present tense is a mistake frequently made by learners and foreigners, and it is also standard in several dialects of English, for instance in the East of England (Trudgill, 2001). This violation was thus much more surprising than its counterpart with *He*, which might explain the presence of the P3-like components even in Native Speakers. The presence of the early negativity in Advanced but not Intermediate speakers suggests that the process of proceduralisation of the explicit

grammatical knowledge of subject-verb agreement in English was more advanced for the first group of learners. However, it should be noted that for Intermediate learners, a better explicit detection of errors in the grammaticality judgment task, as evidenced by a larger d' , meant a more negative-going effect in the early negativities window – and therefore a more native-like response. The negativity of this effect was also correlated with the Time of Instruction in general, which suggests that this early response might not reflect implicit processing as much as was expected and that the critical factor in the native-likeness of the response in this window is not the stay-abroad experience but rather the degree of explicit instruction. The processes underlying the exhibition of this early negativity may thus not be fully unconscious as is often stated, at least in the two learner groups, and the level of instruction might be the prime factor here in the native-likeness of ERP responses.

5 Conclusion

The results of this study shed an interesting light on the early processing of a structure that should be very easy for French learners to master. Qualitative differences between groups were found only for the more “natural” violation, i.e. the one that occurs most commonly and that learners typically produce (with the pronoun *He*). The more artificial violation in the sense that it is rarely if at all produced by learners, triggered attention-dependent responses in all groups. Although both learner groups had a good declarative knowledge of the rules of subject-verb agreement, only Advanced learners exhibited a native-like early anterior negativity. This native-likeness however does not seem to be the result of their study-abroad experience and of the implicit knowledge they could have acquired directly in these conditions, but rather the result of better proceduralisation of their conscious knowledge acquired with years of explicit instruction.

Acknowledgements

This research was supported by an IUF grant awarded to Dr. Emmanuel Ferragne.

References

- R Core Team. (2014). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing.
- Team Audacity. (2013). Audacity version 2.0.3.
- Andringa, S., & Rebuschat, P. (2015). New directions in the study of implicit and explicit learning. *Studies in Second Language Acquisition*, 37(Special Issue 02), 185–196. <http://doi.org/10.1017/S027226311500008X>
- Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., ... Green, P. (2015). lme4: Linear mixed-effects models using ‘Eigen’ and S4: Version 1.1-9.
- Boersma, P., & Weenink, D. (2007). Praat: doing phonetics by computer (Version 5.1.05).
- Bond, K., Gabriele, A., Fiorentino, R., & Alemán Bañon, J. (2011). Individual Differences and the Role of the L1 in L2 Processing: An ERP Investigation. In D. Tanner & J. Herschensohn (Eds.), *Proceedings of the 11th Generative Approaches to Second Language Acquisition Conference (GASLA 2011)* (pp. 17–29). Somerville, MA: Cascadilla Proceedings Project.
- Carhart, R., & Jerger, J. (1959). Preferred method for clinical determination of pure-tone thresholds. *Journal of Speech & Hearing Disorders*, 24(November), 330–345. <http://doi.org/10.1044/jshd.2404.330>

- Chen, L., Shu, H., Liu, Y., Zhao, J., & Li, P. (2007). ERP signatures of subject–verb agreement in L2 learning. *Bilingualism: Language and Cognition*, *10*(2), 161–174. <http://doi.org/10.1017/S136672890700291X>
- Davies, M. (2008). The Corpus of Contemporary American English: 450 million words, 1990–present.
- DeKeyser, R. (2007). Skill acquisition theory. In B. VanPatten & J. N. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 97–113). Mahwah, NJ: Erlbaum.
- Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics. *Journal of Neuroscience Methods*, *134*, 9–21. <http://doi.org/http://doi.org/10.1016/j.jneumeth.2003.10.009>
- Ellis, N. C. (2005). The Handbook of Second Language Acquisition. In C. J. Doughty & M. H. Long (Eds.), *The Handbook of Second Language Acquisition* (pp. 63–103). Oxford, UK: Blackwell Publishing Ltd. <http://doi.org/10.1111/b.9781405132817.2005.00005>.
- Ellis, R. (1994). A theory of instructed second language acquisition. In N. C. Ellis (Ed.), *Implicit and Explicit Learning of Languages* (pp. 79–114). London: Academic Press.
- Fabiani, M., Gratton, G., & Federmeier, K. D. (2007). Event-Related Brain Potentials: Methods, Theory, and Applications. In *Handbook of Psychophysiology* (pp. 85–119). <http://doi.org/http://dx.doi.org/10.1017/CBO9780511546396.004>
- Ferragne, E., Flavier, S., & Fressard, C. (2012). ROCme! Recording of Oral Corpora Made Easy, version 2.1.04. Available at <http://www.ddl.ish-lyon.cnrs.fr/Download/Index.asp?Langue=FR&Page=ROCme>
- Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, *6*(2), 78–84. <http://doi.org/S1364661300018398>
- Gullberg, M., & Indefrey, P. (2003). Language Background Questionnaire. Developed in The Dynamics of Multilingual Processing. Nijmegen, Max Planck Institute for Psycholinguistics.
- Hahne, A., & Friederici, A. D. (1999). Electrophysiological Evidence for Two Steps in Syntactic Analysis: Early Automatic and Late Controlled Processes. *Journal of Cognitive Neuroscience*, *11*(2), 194–205. <http://doi.org/10.1162/089892999563328>
- Hahne, A., & Friederici, A. D. (2001). Processing a second language: late learners' comprehension mechanisms as revealed by event-related brain potentials. *Bilingualism: Language and Cognition*, *4*(2), 123–141. <http://doi.org/10.1017/S1366728901000232>
- Kaan, E., Harris, A., Gibson, E., & Holcomb, P. (2000). The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes*, *15*(2), 159–201. <http://doi.org/10.1080/016909600386084>
- Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Pergamon.
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2015). ImerTest: Tests in Linear Mixed Effects Models: Version 2.0-33.
- Lenth, R. (2016). Ismeans: Least-Squares Means: Version 2.25. Comprehensive R Archive Network (CRAN).
- Lopez-Calderón, J., & Luck, S. J. (2014). ERPLAB: an opensource toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience*, *8*, 213. <http://doi.org/10.3389/fnhum.2014.00213>
- Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing Language Profiles in Bilinguals and Multilinguals. *Journal of Speech, Language, and Hearing Research*, *50*(4), 940–967. [http://doi.org/10.1044/1092-4388\(2007\)067](http://doi.org/10.1044/1092-4388(2007)067)
- Meulman, N., Stowe, L. A., Sprenger, S. A., Bresser, M., & Schmid, M. S. (2014). An ERP study on L2 syntax processing: When do learners fail? *Frontiers in Psychology*,

- 5(SEP). <http://doi.org/10.3389/fpsyg.2014.01072>
- Mueller, J. L. (2005). Electrophysiological correlates of second language processing. *Second Language Research*, 21(2), 152–174. <http://doi.org/10.1191/0267658305sr2560a>
- Neurobehavioral Systems. (2012). *Presentation*.
- Ojima, S., Nakata, H., & Kakigi, R. (2005). An ERP study of second language learning after childhood: effects of proficiency. *Journal of Cognitive Neuroscience*, 17(8), 1212–1228. <http://doi.org/10.1162/0898929055002436>
- Oldfield, R. C. (1971). The Assessment and Analysis of Handedness: The Edinburgh Inventory. *Neuropsychologia*, 9(1), 97–113. [http://doi.org/doi:10.1016/0028-3932\(71\)90067-4](http://doi.org/doi:10.1016/0028-3932(71)90067-4)
- Ping Li, S. S. (2006). Language history questionnaire: A web-based interface for bilingual research. *Behavior Research Methods*, 38(2), 202–10. <http://doi.org/10.3758/BF03192770>
- Pulvermüller, F., Shtyrov, Y., Hasting, A. S., & Carlyon, R. P. (2008). Syntax as a reflex: Neurophysiological evidence for early automaticity of grammatical processing. *Brain and Language*, 104(3), 244–253. <http://doi.org/10.1016/j.bandl.2007.05.002>
- Rebuschat, P., & Williams, J. N. (2013). Implicit learning in second language acquisition. In C. A. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics*. Blackwell Publishing Ltd. <http://doi.org/10.1002/9781405198431>
- Rossi, S., Gugler, M. F., Friederici, A. D., & Hahne, A. (2006). The Impact of Proficiency on Syntactic Second-language Processing of German and Italian: Evidence from Event-related Potentials. *Journal of Cognitive Neuroscience*, 18(12), 2030–2048. <http://doi.org/10.1162/jocn.2006.18.12.2030>
- Sabourin, L., & Stowe, L. A. (2008). Second language processing: when are first and second languages processed similarly? *Second Language Research*, 24(3), 397–430. <http://doi.org/10.1177/0267658308090186>
- Segalowitz, N. B. J. B. N.-C. M. (2004). A Comparison of Spanish Second Language Acquisition in Two Different Learning Contexts: Study Abroad and the Domestic Classroom. *Frontiers: The Interdisciplinary Journal of Study Abroad*, 10, 1–18.
- Tanner, D., Inoue, K., & Osterhout, L. (2013). Brain-based individual differences in online L2 grammatical comprehension. *Bilingualism: Language and Cognition, FirstView*, 1–17. <http://doi.org/10.1017/S1366728913000370>
- Tanner, D., McLaughlin, J., Herschensohn, J., & Osterhout, L. (2013). Individual differences reveal stages of L2 grammatical acquisition: ERP evidence. *Bilingualism: Language and Cognition*, 16(Special Issue 02), 367–382. <http://doi.org/10.1017/S1366728912000302>
- Tanner, J., Wilson, P., & Swets, J. A. (1954). A decision-making theory of visual detection. *Psychological Review*, 61(6), 401–409.
- The Mathworks. (2008). MATLAB and statistics toolbox. Massachusetts, USA: INC. Natick.
- Tokowicz, N., & MacWhinney, B. (2005). Implicit and Explicit Measures of Sensitivity to Violations in Second Language Grammar: An Event-Related Potential Investigation. *Studies in Second Language Acquisition*, 27(2), 173–204. <http://doi.org/10.1017/S0272263105050102>
- Tremblay, A., & Ransijn, J. (2015). LMERConvenienceFunctions: Model Selection and Post-hoc Analysis for (G)LMER Models.
- Trudgill, P. (2001). Modern East Anglia as a dialect area. In J. Fisiak & P. Trudgill (Eds.), *East Anglian English* (pp. 1–12). Cambridge: D.S. Brewer.
- Ullman, M. T. (2001). A neurocognitive perspective on language: The declarative/procedural model. *Nature Reviews Neuroscience*, 2(10), 717–726.
- van Hell, J. G., & Tokowicz, N. (2010). Event-related brain potentials and second language learning: syntactic processing in late L2 learners at different L2 proficiency levels.

Second Language Research, 26(1), 43–74. <http://doi.org/10.1177/0267658309337637>
Weber-Fox, C., & Neville, H. (1996). Maturational constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. *Journal of Cognitive Neuroscience*, 8, 231–256.