Snapshots of Grammaticalization: Differential Eletrophysiological Responses to Grammatical Anomalies with Increasing L2 Exposure

Darren Tanner, Lee Osterhout, and Julia Herschensohn
University of Washington

1. Introduction

It is frequently observed that second language (L2) learners show significant problems producing inflectional morphology. For example, many researchers have shown that L2 learners in the early stages of acquisition may show no function words or morphemes in their speech (Klein & Perdue, 1997; Vainikka & Young-Scholten, 1996) and that morphological errors may persist even in advanced learners who otherwise have a strong command of the L2's syntax (Herschensohn, 2001; Lardiere, 2007; White, 2003). While it is clear that acquiring and producing L2 inflectional morphology may present a formidable task to learners, much less is known about how and at what rate learners integrate L2 morphological knowledge into their real-time comprehension and processing system. Furthermore, little is known about the stages learners pass though during acquisition of L2 inflectional morphology. This study therefore investigates the processing of agreement morphology in the early stages of L2 acquisition using event-related potentials (ERPs). In particular, we focus on the processing of subject-verb agreement in beginning and intermediate native English speakers acquiring German as an L2.

Subject-verb agreement is a feature present in many Indo-European languages, including both English (*I drink beer* vs. *He drinks beer*) and German (*Ich trinke Bier* vs. *Er trinkt Bier*). The grammatical features underlying knowledge of such agreement in native speakers is fairly well understood (see e.g., Corbett, 2006) and processing of agreement has been actively studied in the native language speech production (e.g., Bock & Miller, 1991; Vigliocco & Nicol, 1998) and comprehension literature (e.g., Nicol, Forster, & Veres, 1997; Osterhout & Mobley, 1995; Pearlmutter, Garnsey, & Bock, 1999). With regard to processing agreement in comprehension, behavioral measures of processing such as timed reading and eye tracking have shown that agreement mismatches, such as **The key to the cabinets <u>were</u> rusty* elicit longer eye fixation and reading times and more regressive eye movements at the offending verb than agreement matches, even when the subjects were asked simply to read for comprehension

[*] We received financial support from Grant Number R01DC01947, funded by the NIDCD. We would also like to thank the participants, Judy McLaughlin, and the other members of the Neurocognition of Language Lab at the University of Washington.

(Nicol et al., 1997; Pearlmutter et al., 1999). The grammatical features of the subject, such as person and number, therefore seem to be automatically available and integrated upon encountering the verb during sentence comprehension. This is not a trivial finding, given that agreement morphology on verbs provides largely redundant information (that is, the grammatical features of the subject are already specified on the subject) and that the morphology itself is devoid of semantic content. Moreover it has been shown in off-line sentence interpretation tasks in English, which has a relatively impoverished agreement system, that agreement information is largely ignored by native speakers in interpreting and identifying grammatical relations within a sentence (MacWhinney, Bates, & Kliegl, 1984). Nonetheless, agreement mismatches still seem to cause English speakers on-line processing difficulty, even when computing agreement is irrelevant to the experimental task.

There is also a large body of evidence from ERP studies that processing of agreement mismatches elicits neural responses distinct from those elicited by correctly inflected words. In particular it has been shown that, relative to grammatical control words, incorrectly inflected verbs (as in **The elected officials hopes to succeed*) reliably elicit a large positive-going wave with an onset around 500ms post-stimulus which lasts for several hundred milliseconds and which is largest over posterior regions of the scalp (Osterhout & Mobley, 1995). This effect, known as the P600, has been found in response to agreement mismatches and other syntactic anomalies in a number of languages and is therefore taken to be a reliable marker of syntactic processing (Osterhout & Holcomb, 1995). Some studies of syntactic processing have reported an additional negative-going wave preceding the P600 with an onset of between 100-400ms post-stimulus, sometimes with a more left anterior distribution (the (early) left anterior negativity, (E)LAN) (Hahne & Friederici, 1999; Neville, Nicol, Barss, Forster, & Garrett, 1991; Osterhout & Mobley, 1995).

The P600 and LAN effects contrast with another well-studied ERP effect related to language processing, the N400. The N400 is a more negative going wave, largest over centro-parietal scalp regions (sometimes with a larger right hemisphere distribution), that usually peaks around 400ms after stimulus presentation. The N400 is sensitive to aspects of lexical processing, such as word/nonword status, word frequency and a word's semantic relatedness to a preceding context (Kutas & Van Petten, 1994). Although the exact cognitive processes underlying the N400 and P600 remain unknown, they reliably correlate with lexical and syntactic processing, respectively. Important for the current experiment is that processing of agreement anomalies reliably correlates with a P600 effect in native speakers of a language, indicating that agreement processes during comprehension are syntactic in nature.

Recently there has been increased interest in using ERPs to study L2 syntactic processing (see Mueller, 2005 for a review). A few patterns of L2 processing emerge from these studies. It seems that factors such as age of acquisition, proficiency, and L1-L2 similarity interact in complex ways in shaping patterns of neural responses to syntactic processing in L2 learners. In particular, several studies have shown qualitatively similar responses in L1 and L2 processing, though with the L2 learners sometimes showing delayed and reduced neural responses. Other studies have reported a lack of ERP effects, particularly in very late learners and with grammatical structures that differ between the L1 and L2. The question of proficiency, however, is less clear. Rossi and colleagues (Rossi, Gugler, Friederici, & Hahne, 2006) compared low and high proficiency L2 learners and found that while high proficiency learners of German and Italian both showed biphasic LAN + P600 responses to subject-verb agreement violations, low proficiency learners showed only a P600 with a smaller amplitude and later onset compared with high proficiency learners. That is, the P600 seen in the low proficiency group was qualitatively similar to that found in the more advanced group, but showed quantitative differences. However, their low proficiency learners were hardly beginners. The L2 German and Italian learners had on average 3.5 and 2.1 years of classroom instruction, respectively, and self-rated their proficiency as intermediate (approximately 3 on a 6-point scale). Tokowicz and MacWhinney (2005) studied low proficiency classroom-instructed Anglophone learners of Spanish. Their subjects showed a P600 to verb tense and gender agreement violations, but no response to determiner-noun number agreement violations. Thus, their learners showed a P600 to syntactic anomalies, but it was modulated by L1-L2 grammatical similarity. However, they averaged over students enrolled in each of the first four semesters of university Spanish courses, thus obscuring any differences in brain response between the beginning and more intermediate learners.

It may be the case that learners use different cognitive resources in processing morphosyntax in the very earliest stages of L2 acquisition versus in intermediate and later stages. Unfortunately traditional measures of processing, such as timed reading or eye tracking, are not fully sensitive to qualitative aspects of processing. They indicate whether or not a subject is experiencing processing difficulty, but not the nature of that difficulty. ERPs, however, provide an ideal tool to investigate qualitative aspects of processing, given that they are sensitive to the difference between syntactic and lexical processing, indexed by the P600 and N400, respectively. The data reported below are part of a cross-sectional study of agreement processing in low and intermediate proficiency L2 learners of German whose native language is English. In particular we investigate the processing of subject-verb agreement anomalies using ERPs in order to investigate the nature of processing in the earlier stages of L2 acquisition and to assess changes in processing signatures associated with increasing exposure and proficiency.

2. Methods

Participants. Our participants included 13 native speakers of German, 13 native English speakers enrolled in third year German courses at the University of Washington (300-level courses), and 20 native English speakers enrolled in the final course of the first year German sequence (German 103). In order to more accurately assess differences in learning speed, first year subjects were then divided into two groups (quick learners, n=10, and slow learners, n=10) based on behavioral data collected during the experiment (grammaticality judgments). The criterion for grouping was a median split on d´ scores (a measure of sensitivity to ungrammaticalities).

Materials. Sixty grammatical-ungrammatical sentence pairs were constructed. Each sentence began with the first person singular pronoun *Ich* and was followed by the critical verb, as in (1).

(1) Ich <u>trinke/*trinkt</u> nie Alkohol.
I drink/*drinks never alcohol
'I never drink alcohol.'

Grammaticality was manipulated by the agreement features on the verb; verbs either agreed with the subject pronoun (grammatical) or were marked with third person singular agreement (ungrammatical). The first person pronoun/third person verb contrast was chosen because this grammatical contrast is also expressed in English. This will henceforth be referred to as the verbal agreement condition. All lexical items were chosen from the first seven chapters of the textbook used in first year German courses at the University of Washington in order to ensure that learners were familiar with the words used in the experiment. The sentences were distributed across two experimental lists such that each subject saw only one version of each sentence and such that each subject saw thirty grammatical and thirty ungrammatical sentences.

Each list also contained 140 other sentences, 70 of which were ungrammatical. Sixty of the remaining ungrammatical sentences contained number disagreement between a determiner or quantifier and a noun (2a-b) (the noun agreement condition) [1]; ten filler contained an extra auxiliary verb (2c).[2]

(2)

a. Viele/*ein <u>Bücher</u> liegen auf dem Tisch.
Many/*a books lies on the table
'Many books are on the table.'

b. Dein/*viele <u>Photo</u> ist sehr schön.
Your/*many photo is very beautiful
'Your photo is very beautiful.'

c. *Mein Bruder macht <u>sind</u> seine Arbeit.
*My brother makes are his work
'My brother is doing (are) his work.'

Each list was pseudo-randomized. In all, each list contained 200 sentences, half of which contained an ungrammaticality.

---

[1] ERP results from this condition will not be discussed here.
[2] Behavioral and ERP data from the filler sentences were not analyzed.

Procedure. Participants were tested in a single session and were randomly assigned to one of the stimulus lists. Participants were seated in a comfortable recliner in front of a CRT monitor. Each trial consisted of a blank screen for 1000ms, followed by a fixation cross, followed by a stimulus sentence, presented one word at a time. The fixation cross and each word appeared on the screen for 475ms followed by a 250ms blank screen between words. Sentence-ending words appeared with a full stop followed by a response prompt asking subjects if the sentence was "good" or "bad". Subjects were instructed to respond "good" if they felt it was a well-formed, grammatical sentence in German and "bad" if they felt it was ungrammatical or violated some rule of German. Subjects were also informed to guess if they were unsure about a sentence's grammaticality. Additionally, two versions of each list were created in order to counterbalance the "good" response hand, resulting in four overall lists (two lists, each with a left and right hand "good" response version).

Data acquisition and analysis. Continuous EEG was recorded from 19 scalp sites and averaged offline. Data from three midline electrodes, Fz (frontal), Cz (central), and Pz (posterior), will be reported here, as the effects of interest are strongest over midline electrodes. Electrodes were referenced to an electrode placed over the left mastoid and were amplified with a bandpass of 0.01-100Hz (3db cutoff) by an SAI bioamplifier system; an additional offline low-pass filter of 15Hz was applied prior to data analysis. ERPs, time-locked to the onset of the critical word (the words underlined in 1 and 2, above), were averaged off-line for each subject at each electrode site, relative to a 100ms prestimulus baseline.

Behavioral results were quantified using d′ scores, calculated for each participant, where d′ = z(Hit) – z(False Alarm) (i.e., the difference between the standardized probabilities of correctly identifying a grammatical sentence as grammatical and incorrectly identifying an ungrammatical sentence as grammatical); a d′ of 0 means chance performance and a d′ of 4 means nearly perfect performance. ERP components of interest were quantified by computer as mean voltage within a window of activity. In accordance with previous literature and visual inspection of the data, the following time windows were chosen: 300-500ms (N400) and 500-800ms (P600). Within each time window ANOVAs were calculated with Group (Native, Third, First(Quick), First(Slow)) as a between-subjects factor and Grammaticality (Grammatical, Ungrammatical) and Electrode (Fz, Cz, Pz) as repeated measures factors. The Greenhouse-Geisser correction for inhomogeneity of variance was applied to all repeated measures with greater than one degree of freedom in the numerator. In such cases, the corrected *p* value is reported. Significant or near-significant interactions with Group were followed up with separate repeated measures ANOVAs for each group.

3. Results

Behavioral results. Means and standard deviations for the grammaticality judgment task are presented in Table 1.

Table 1. Means and standard deviations of d′ scores

	Verb Agreement		Noun Agreement	
Group	Mean	Std. Dev.	Mean	Std. Dev
First (slow)	2.43	1.08	0.61	0.57
First (quick)	3.97	1.13	2.14	0.54
Third	3.61	1.40	3.01	0.95
Native	4.85	0.93	4.42	1.38
Total	3.78	1.41	2.70	1.67

An ANOVA with Structure (Verb or Noun) as a within-subjects factor and Group as a between-subjects factor showed a main effect of Structure, $F(1, 42) = 26{,}929$, $p < .001$, indicating that participants were overall more sensitive to grammaticality in the verbal agreement condition than in the nominal agreement condition. There was a main effect of Group, $F(3, 42) = 32.969$, $p < .001$, and a significant Structure by Group interaction, $F(3, 42) = 2.829$, $p = .050$. This interaction results from the fact that the two first year groups were markedly less sensitive to grammaticality in the nominal versus the verbal agreement condition while the third year and native groups were only marginally so. The main effect of group was followed up with a Tukey's post hoc test, which showed no significant difference in overall mean between the first year "quick" learners and third year group ($p = .852$); all other pairings were significantly different (all *p*s $< .001$).

ERP Results. Grand average waveforms for the three midline electrodes are presented for each group in Figure 1.

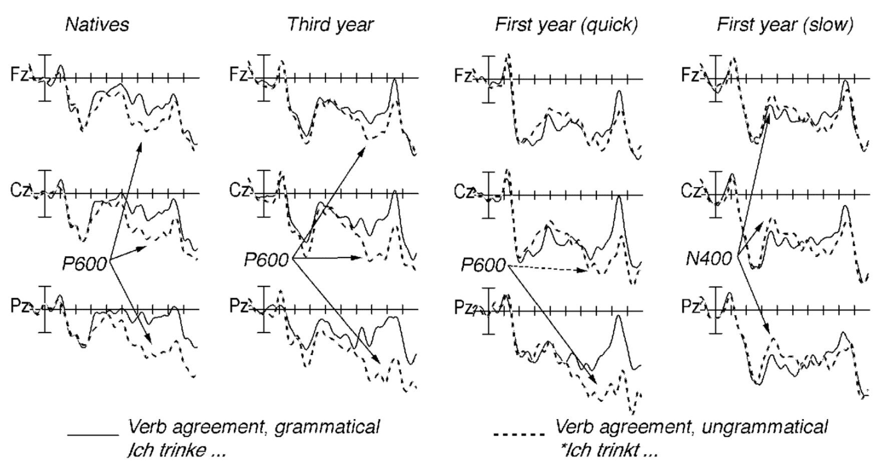

Figure 1. Grand average waveforms for midline electrodes (Fz, Cz, Pz) for native speakers (n=13), third year learners (n=13), first year (quick) learners (n=10), and first year (slow) learners (n=10) to agreeing (solid line) and disagreeing (dashed line) verbs.

Statistical analyses show no significant main effect of Grammaticality or Grammaticality x Electrode interaction in the 300-500ms time window (Fs < 1); however there was a significant Group x Grammaticality interaction, $F(3, 42) = 3.944$, $p = .014$, indicating that responses differed by group. To follow up on this, separate ANOVAs were calculated for each group. Native speakers showed a trend toward a positive deflection in brain responses to ungrammaticalities (onset of a P600), $F(1, 12) = 3.524$, $p = .085$. There were no effects in the third year and first year (quick) groups, Fs < 2.7, *p*s > .123. The first year (slow) group, however, showed a significant N400, $F(1, 9) = 8.713$, $p = .016$, that was equal over electrodes.

In the 500-800ms time window, responses were overall more positive-going in response to ungrammaticalities, $F(1, 42) = 28.028$, $p < .001$, and this response was strongest over posterior electrodes, $F(2, 84) = 6.593$, $p = .003$ (a P600). However, there was a near-significant Group x Grammaticality x Electrode interaction, $F(6, 84) = 2.015$, $p = .083$, indicating that this effect differed across groups; separate group ANOVAs followed up on this effect. Native speakers showed a significant P600, $F(1, 12) = 25.960$, $p < .001$, that was strongest over posterior electrodes, $F(2, 24) = 3.818$, $p = .043$. Third year learners showed a broadly distributed P600, $F(1, 12) = 17.532$, $p = .001$, but there was no difference across electrodes. First year (quick) learners showed a P600, but this effect was restricted to posterior electrodes, $F(2, 18) = 9.551$, $p = .003$; first year (slow) learners showed no effects in this time window.

In order to assess the relationship between behavioral sensitivity to subject-verb agreement and brain responses, differences between participants' mean amplitude in the 500-800ms time window to grammatical and ungrammatical verbs (i.e., the size of individuals' P600) were regressed onto their d′ scores to the verb agreement condition. Mean amplitudes were taken at electrode Pz where the P600 effects were strongest. The correlation between d′ score and P600 for all first year learners combined was significant, $r = .532$, $p = .016$, for third year learners the correlation neared significance, $r = .501$, $p = .081$, and for all learners (first and third year) combined the correlation was highly significant, $r = .533$, $p = .001$, Figure 2. Thus, there was a linear relationship between performance on the grammaticality judgment task and brain responses for learners, with larger P600s seen in those who showed greater sensitivity to agreement mismatches. For native speakers there was no such relationship between d' and P600 amplitude, $r = .177$, $p = .563$.

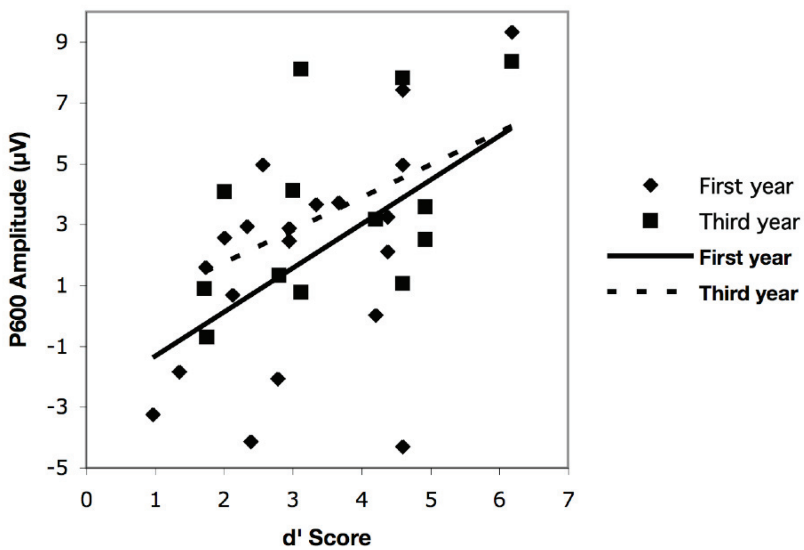

Figure 2. All learners' P600 amplitudes over Pz by d´ score. Separate regression lines are provided for first year (n=20) and third year (n=13) participants.

4. Discussion

Results from this study can be briefly summarized as follows. All groups showed behavioral sensitivity to subject-verb agreement in German, with better than chance performance on the grammaticality judgment task, though the first year (slow) group performed significantly worse than the more advanced groups. All groups also showed neural sensitivity to agreement violations, though the responses differed between groups. Native speakers, third year learners, and first year (quick) learners all showed a P600 to subject-verb agreement violations, though the P600 showed a slightly later onset in the learner groups and was restricted to posterior electrodes in the first (quick) group. The first (slow) group showed an N400 to disagreeing verbs. This qualitatively different response seen in the lowest proficiency group suggests that despite the learners' sensitivity to verbal agreement in German, different cognitive resources underlie its processing.

This qualitative difference is particularly relevant to theories of the acquisition of L2 morphosyntax. For example, many scholars interested in the nature of L2 grammatical representations have posited that learners will initially transfer their first language (L1) grammar to their L2 (e.g., Schwartz & Sprouse, 1996; White, Valenzuela, Kozlowska-MacGregor, & Leung, 2004). On this approach any grammatical feature or rule present in a speaker's L1 should be present in the learner's interlanguage grammar from the initial state of acquisition. Processing approaches to L2 acquisition have also argued for strong transfer of L1 processing routines, which should form the basis of the developing L2 systems (e.g., MacWhinney, 2005; Sabourin & Haverkort, 2003; Sabourin & Stowe, 2008). For example, the Competition Model predicts that "whatever can transfer will" (MacWhinney, 2005, p. 55) and assumes that the L2 lexicon, phonology, and grammar share neural resources with those of the L1. Processing systems will therefore be 'parasitic' on the constructs supplied by the L1 (MacWhinney, 1997, p. 119), and processing signatures for grammatical phenomena that are alike in the L1 and L2 will be qualitatively similar, eliciting a P600 (Tokowicz & MacWhinney, 2005).

Others hold that learners bootstrap rules from learned formulae, inducing abstract rules after extracting regularities from rote-learned chunks (Myles, Hooper, & Mitchell, 1998; Wong Fillmore, 1976; Wray, 2002; Zobl, 1998). If this is the case, one might expect that beginning learners of German may be sensitive to verbal agreement because of grammatical transfer from English, but that they may initially memorize inflected verb forms such as *trinkt* ('drink'.3rd.sg) as unanalyzed wholes, rather than stem (*trink-*) plus affix (*-t*) sequences. Given a first person singular subject (*Ich*), learners may then perceive *trinkt* as an unexpected word form rather than as a violation of a grammatical rule; processing signatures would therefore be indicative of lexical, rather than syntactic, processing in the initial stages of L2 acquisition. Learners might later induce a productive morphosyntactic rule and decompose these verbal 'chunks' into stem plus affix sequences. At this point learners would have grammaticalized the inflectional rule and show signatures of syntactic processing. The current data are consistent with this second model.

There is another possible explanation, however. Namely, the behavioral/brain response correlations seen in the data may not show a developmental continuum, but rather be a result of different learning strategies. For example, it could be the case that memorizing subject-verb sequences is a less effective learning strategy than identifying a discontinuous dependency between the features of the subject and verb morphology. Students employing the memorization strategy (i.e, those who show an N400) may be less sensitive to ungrammaticalities involving agreement morphology, resulting in lower d′ scores; students employing the second, more effective, strategy (i.e., those showing a P600) would show greater sensitivity and higher d′ scores. Given the current data there is no clear way to choose between this possibility and the previously discussed more developmental alternative.

The current data, however, bear a striking resemblance to data reported by Osterhout and colleagues (2006). In a longitudinal study of novice L2 French learners, they found that subjects who did well judging sentences (median split on d′ scores, similar to the current study) showed an N400 effect to subject-verb agreement violations after just one month of exposure to French. After four months the N400 had been replaced by a P600, which was native-like by the end of the first year of instruction. Thus, the first (slow) learners in the current study look very similar to Osterhout's subjects during their first testing session, while the more advanced learners in the current study resemble Osterhout's learners at the end of the study. It seems highly likely that the current study captures the same transition from N400 to P600 cross-sectionally that Osterhout and colleagues found longitudinally.

A final point is that the N400-P600 discontinuity seen in the current study and in the Osterhout study was not seen in other studies investigating low-proficiency learners (Rossi et al., 2006; Tokowicz & MacWhinney, 2005). Rossi studied low proficiency learners of L2 German processing subject-verb agreement violations and found a P600 effect; no N400 was present in her data. However, as mentioned above, the low-proficiency learners in her study were likely much more advanced than the first year subjects in the current experiment. It may have been the case that learners in her study had already passed through the N400 phase. Alternately, the N400 found in the current experiment may relate to distinct learning environments experienced by participants in the two studies. Classroom learning, as in this experiment, may lend itself more readily to rote memorization of verb forms in the early stages of acquisition; Rossi's subjects, on the other hand, reported a mix of both classroom learning and some immersion in an L2 environment. It is unclear what effect this may have had on ERP results. As for Tokowicz and MacWhinney's study, they base their claim that processing mechanisms for similar structures are shared across the L1 and L2 on the fact that their subjects showed a P600 to violations of verb tense in Spanish. However, as previously discussed, Tokowicz and MacWhinney averaged across subjects in the first four semesters of university Spanish; thus, any differences in the most beginning learners and more advanced learners would be lost in the final statistics. It may also be the case that the lack of N400 seen in their data relates to their choice of verb tense (e.g., *Su abuela cocina/*cocinando muy bien* "His grandmother cooks/*cooking very well"), and not subject-verb agreement, as the relevant contrast. A future avenue for research should be to contrast agreement and tense violations within learner groups at different proficiencies.

5. Conclusion

This study investigated the processing of subject-verb agreement anomalies in native speakers of German and beginning and intermediate Anglophone L2 learners of German. Results showed that L2 learners were sensitive to grammatical errors in the L2, indexed by both behavioral and electrophysiological responses. Learners enrolled in third year classes showed qualitatively similar brain responses to those found in native speakers, indicating that learners can use similar neurocognitive resources in processing L2 agreement, at least when the grammatical contrast is similar across the two languages. Learners enrolled in first year courses who performed well judging sentence grammaticality showed a P600, though with a later onset and smaller scalp distribution relative to native speakers; first year learners who performed poorer (but still better than chance) at judging sentence grammaticality showed a small N400 effect. These results corroborate previous longitudinal findings from L2 French. We take the current findings to indicate that learners pass through discontinuous stages during the acquisition of L2 agreement morphology: learners process inflected verbs as unanalyzed units in the early stages of L2 acquisition, subsequently decomposing them into stem plus affix sequences and inducing a productive morphosyntactic rule.

References

Bock, Kathryn, & Miller, C. A. (1991). Broken Agreement. *Cognitive Psychology,* 23, 45-97.

Corbett, Greville. (2006). *Agreement*. Cambridge: Cambridge University Press.

Hahne, Anja, & Friederici, Angela D. (1999). Eletrophysiological evidence for two steps in syntactic analysis: Early automatic and late controlled processes. *Journal of Cognitive Neuroscience,* 11(2), 194-205.

Herschensohn, Julia. (2001). Missing inflection in second language French: accidental infinitives and other verbal deficits. *Second Language Research,* 17(3), 273-305.

Klein, Wolfgang, & Perdue, Clive. (1997). The Basic Variety (or: Couldn't natural languages be much simpler?). *Second Language Research,* 13(4), 301-347.

Kutas, Marta, & Van Petten, Cyma K. (1994). Psycholinguistics electrified: Event-related brain potential investigations. In M. A. Gernsbacher (Ed.), *Handbook of Psycholinguistics* (pp. 83-143). San Diego: Academic Press.

Lardiere, Donna. (2007). *Ultimate attainment in second language acquisition: A case study*. Mahwah, NJ: Erlbaum.

MacWhinney, Brian. (1997). Second language acquisition and the Competition Model. In A. M. B. De Groot & J. F. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives* (pp. 113-142). Mahwah, NJ: Lawrence Erlbaum.

MacWhinney, Brian. (2005). A unified model of language acquisition. In J. F. Kroll & A. M. B. De Groot (Eds.), *Hanbook of bilingualism: Psycholinguistic approaches* (pp. 49-67). Oxford: Oxford University Press.

MacWhinney, Brian, Bates, Elizabeth, & Kliegl, R. (1984). Cue validity and sentence intepretation in English, German, and Italian. *Journal of Verbal Learning and Verbal Behavior,* 23, 127-150.

Mueller, Jutta L. (2005). Electrophysiological correlates of second language processing. *Second Language Research,* 21(2), 152-174.

Myles, Florence, Hooper, Janet, & Mitchell, Rosamond. (1998). Rote or Rule? Exploring the Role of Formulaic Language in Classroom Foreign Language Learning. *Language Learning,* 48(3), 323-364.

Neville, Helen J., Nicol, Janet, Barss, Andrew, Forster, Kenneth, & Garrett, Merrill. (1991). Syntactically based sentence processing classes: Evidence from event-related brain potentials. *Journal of Cognitive Neuroscience,* 3(2), 151-165.

Nicol, Janet L., Forster, K. I., & Veres, C. (1997). Subject-Verb Agreement Processes in Comprehension. *Journal of Memory and Language,* 36(4), 569-587.

Osterhout, Lee, & Holcomb, Phillip J. (1995). Event-related brain potentials and language comprehension. In M. D. Rugg & M. G. H. Coles (Eds.), *Electrophysiology of mind: Event-related brain potentials and cognition*). Oxford: Oxford University Press.

Osterhout, Lee, McLaughlin, Judith, Pitkanen, Ilona, Frenck-Mestre, Cheryl, & Molinaro, Nicola. (2006). Novice Learners, Longitudinal Designs, and Event-Related Potentials: A Means for Exploring the Neurocognition of Second Language Processing. *Language Learning,* 56(s1), 199-230.

Osterhout, Lee, & Mobley, Linda. (1995). Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language,* 34, 739-773.

Pearlmutter, Neal J., Garnsey, Susan M., & Bock, Kathryn. (1999). Agreement Processes in Sentence Comprehension. *Journal of Memory and Language,* 41(3), 427-456.

Rossi, Sonja, Gugler, Manfred F., Friederici, Angela D., & Hahne, Anja. (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.

Sabourin, Laura, & Haverkort, Marco. (2003). Neural substrates of representation and processing of a second language. In R. van Hout, A. Hulk, F. Kuiken & R. Towell (Eds.), *The lexicon-syntax interface in second language acquisition* (pp. 175-195). Amsterdam/Philadelphia: John Benjamins.

Sabourin, Laura, & Stowe, Laurie A. (2008). Second language processing: when are first and second languages processed similarly? *Second Language Research,* 24(3), 397-430.

Schwartz, Bonnie D., & Sprouse, Rex. (1996). L2 cognitive states and the full transfer/full access model. *Second Language Research,* 12, 40-72.

Tokowicz, Natasha, & MacWhinney, Brian. (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.

Vainikka, Anne, & Young-Scholten, Martha. (1996). Gradual development of L2 phrase structure. *Second Language Research,* 12, 7-39.

Vigliocco, Gabriella, & Nicol, Janet. (1998). Separating hierarchical relations and word order in language production: is proximity concord syntactic or linear? *Cognition,* 68(1), 13-29.

White, Lydia. (2003). Fossilization in steady state L2 grammars: Persistent problems with inflectional morphology. *Bilingualism: Language and Cognition,* 6(2), 129-141.

White, Lydia, Valenzuela, Elena, Kozlowska-MacGregor, Martyna, & Leung, Yan-Kit Ingrid. (2004). Gender and number agreement in nonnative Spanish. *Applied Psycholinguistics,* 25, 105-133.

Wong Fillmore, Lily. (1976). *The second time around: Cognitive and social strategies in second language acquisition.* Unpublished doctoral dissertation, Stanford University, Stanford, CA.

Wray, Alison. (2002). *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.

Zobl, Helmut. (1998). Representational changes: From listed representations to independent representations of verbal affixes. In M.-L. Beck (Ed.), *Morphology and its interfaces in second language knowedlge* (pp. 339-371). Amsterdam: Benjamins.