Electrophysiological Markers of Interference and Structural Facilitation in Native and Nonnative Agreement Processing

Darren Tanner, Janet Nicol, Julia Herschensohn, and Lee Osterhout

1. Introduction

Studies of second language (L2) learning have long noted that learners acquiring an L2 post-puberty rarely achieve native-like mastery of numerous aspects of the target language, including phonology and morphosyntax (e.g., Flege, Munro, & MacKay, 1995; Johnson & Newport, 1989). One aspect of morphosyntax that has been seen to be especially problematic in both production and comprehension is inflectional morphology, including grammatical agreement (e.g., Hopp, 2010; Jiang, 2004; Lardiere, 2007). One proposed explanation for this difficulty is that native (L1) and L2 processing use qualitatively similar mechanisms, but that L2 processing is slower and less automatic, leading to difficulty retrieving and integrating grammatical information in real-time (Hopp, 2010; Lardiere, 2000). Alternately, it has been suggested that grammatical and morphological processing in late L2 learners is qualitatively different from L1, regardless of L2 proficiency or L1-L2 pairing (Clahsen & Felser, 2006; Clahsen, Felser, Neubauer, Sato, & Silva, 2010). Much of this research has used behavioral measures (e.g., grammaticality judgments, error rates in corpus analyses, or reaction time data), and as such cannot adequately address whether the observed differences between L1 and L2 speakers are indeed qualitative or simply quantitative. Measures of neural activity underlying processing, like event-related brain potentials (ERPs), provide both quantitative and qualitative information about language processing and thus can potentially adjudicate between these competing possibilities. The research presented here therefore uses ERPs to study neural correlates of agreement processing in L1 English and highly proficient, late L1 Spanish-L2 English bilinguals. We focus on the neural substrates of L1 and L2 morphosyntactic processing, as well as morphological and syntactic factors related to processing interference and facilitation.

1 Darren Tanner, Pennsylvania State University; Janet Nicol, University of Arizona; Julia Herschensohn and Lee Osterhout, University of Washington. Please direct correspondence to Darren Tanner at dst12@psu.edu. This research was supported by grant R01DC01947 from the NIDCD to Lee Osterhout and grant BCS-0951595 from the NSF to Lee Osterhout and Darren Tanner. Darren Tanner also received support from the William Orr Dingwall Foundation Neurolinguistics Dissertation Fellowship.
It is important to note that even native speakers of a language do not always show error-free processing of L1 morphosyntax during normal language use. It is well-attested that speakers can show variability in subject-verb number agreement marking when the subject noun phrase (NP) contains an embedded modifier with an NP whose number feature mismatches with that of the head noun (e.g., *The key to the wooden cabinets*…). In such cases, speakers sometimes produce a plural-marked verb (e.g., *were*), even though the embedded plural noun is not a legitimate syntactic licensor of agreement (Francis, 1986). This phenomenon, frequently referred to as *agreement attraction*, has received considerable attention in the experimental psycholinguistics literature. Attraction phenomena show that the subject-verb agreement dependency can be ‘leaky’ and open to influence from non-grammatical information. Moreover, numerous syntactic, semantic, and morphophonological factors have been shown to modulate error rates in verb agreement production in laboratory experiments (e.g., Bock & Miller, 1991; Franck, Vigliocco, & Nicol, 2002; Hartsuiker, Schriefers, Bock, & Kikstra, 2003; Vigliocco, Hartsuiker, Jarema, & Kolk, 1996). It has also been shown that comprehension of subject-verb agreement can be similarly ‘error prone’ when a plural attractor noun (*cabinets* in the above example) intervenes between the head noun (*key*) and verb. The most consistent finding in this regard is that difficulty associated with processing ungrammatical subject-verb agreement (e.g., slower reading times) is mitigated following singular head-plural attractor noun combinations relative to singular-singular combinations (e.g., Pearlmutter, Garnsey, & Bock, 1999; Wagers, Lau, & Phillips, 2009). Some have also found that plural attractor nouns cause difficulty even in grammatical sentences (e.g., *The key to the cabinets was…*; Nicol, Forster, & Veres, 1997; Pearlmutter et al., 1999), though this finding is not consistent across studies (Wagers et al., 2009).

Identifying factors that modulate error rates in language production and comprehension can thus provide an important tool in helping understand the cognitive mechanisms underlying successful language use. If a given manipulation can affect rates of attraction, we can infer that that type of information is relevant to the computation of agreement dependencies. Moreover, if it is the case that L1 and L2 users show similar ERP responses to agreement violations, as well as similar attraction profiles and sensitivity to syntactic structural modulations of these, this would suggest a continuity between L1 and L2 processing, both in terms of use of morphological and structural cues used in establishing syntactic dependencies and the neural systems underlying the processing of those dependencies.

Here we focus on one syntactic manipulation known to modulate attraction error rates in production. Studies of language production have shown that the number of errors induced by a plural attractor is greater when the attractor noun is embedded in a prepositional phrase (PP; e.g., *The key to the wooden cabinets*…) than when the attractor is embedded in a relative clause (RC; e.g., *The key that opened the cabinets*) that is matched for length and semantic relatedness to the head noun (Bock & Cutting, 1992; Solomon & Pearlmutter,
While theoretical explanations of this effect vary (see e.g., Badecker & Kuminiak, 2007; Eberhard, Cutting, & Bock, 2005; Franck et al., 2002), it suggests that attraction is to some extent clause-bound, in that plural attractors exert significantly less interference when there is greater structural separation from the head noun. While the clause-bounding effect is well-attested in language production, no published studies have investigated whether it also surfaces in agreement comprehension. Thus, a first goal of this study is to establish if and how the syntactic complexity of a modifying phrase containing a potential agreement attractor affects responses to agreement violations in native English speakers. Moreover, few studies have used ERPs to study profiles of agreement attraction in native speakers (though see Kaan, 2002; Severens, Jansma, & Hartsuiker, 2008 for findings from Dutch). The current study thus helps provide a baseline for such effects in English.

With regard to L2 processing, using ERPs to study attraction interference and structural modulation of attraction effects can provide evidence regarding the aforementioned claims of a qualitative difference between L1 and L2 processing. For example, Clahsen and colleagues’ Shallow Structure Hypothesis (SSH: Clahsen & Felser, 2006) holds that L2 sentence parsing involves only shallow syntactic structures, in that learners (unlike L1 users) underutilize structural information when establishing long-distance dependencies within sentences. Learners instead rely more on linear, thematic, and pragmatic cues to sentence meaning and structure. How might shallow parsing impact profiles of attraction interference? One possibility is that learners will adopt default strategies for agreement based on thematic relations, perhaps assuming that the first NP encountered (i.e., the head noun) is always the agent, and the agent thematically controls agreement via semantic mechanisms. On this scenario there should be no interference from plural attractor nouns when processing agreement on the verb. Alternately, learners may rely strictly on linear relations within the sentence. For example, learners may assume that the first NP (regardless of thematic role) always controls agreement. This would result in a similar outcome to the default thematic strategy, though for different reasons. Alternately, learners might compute agreement dependencies strictly using a local linear adjacency heuristic, where the most proximal noun to the verb (in this case the attractor noun) would always control agreement. If this is the case, L2 learners would perceive grammatical sentences with singular head and plural attractor nouns (e.g., *The key to the wooden cabinets is…*) as ungrammatical because of the agreement mismatch in the local string *cabinets is*. Importantly, if L2 learners are insensitive to syntactic structural manipulations, there should be no differences in brain responses to agreement violations based on the syntactic structure of the subject modifier (PP versus RC).

In assessing L1 and L2 English speakers’ sensitivity to agreement attraction and structural complexity of the subject NP we focus specifically on one ERP component known to be sensitive to syntactic manipulations. The P600 is a large positive-going brainwave with an onset around 500ms after stimulus onset, and is elicited by a range of sentence-embedded morphosyntactic anomalies (e.g.,
Friederici, Hahne, & Mecklinger, 1996; Gouvea, Phillips, Kazanina, & Poeppel, 2010; Osterhout & Mobley, 1995). Importantly, the amplitude of the P600 effect has been shown to correlate with the severity of the anomaly being processed (e.g., Gouvea et al., 2010; Nevins, Dillon, Malhotra, & Phillips, 2007; Osterhout, Holcomb, & Swinney, 1994). Thus, relative size of the P600 effect may index the degree of attraction interference (cf. Kaan, 2002). ERP studies of L2 processing generally find qualitatively similar neural responses to morphosyntactic violations (i.e., P600s), provided there is similar rule overlap between the L1 and L2 (e.g., Foucart & Frenc-Mestre, 2011; Tokowicz & MacWhinney, 2005). Thus, we expect subject-verb agreement violations to elicit P600 effects in L1 Spanish-L2 English bilinguals, as this type of agreement is shared by both languages. The crucial questions investigated here are whether the structure of the subject NP modulates attraction interference and whether, compared to monolingual English speakers, the bilingual group shows qualitatively similar profiles of attraction interference and modulation by subject NP syntactic structure.

2. Method

Participants. Our participants included 31 monolingual native English speakers and 20 L1 Spanish-L2 English late bilinguals. The bilinguals grew up with no English in the home, and were late arrivals in an English-speaking environment (mean age of arrival = 23.9 years, SD = 6.4, range: 15-40) and had all been immersed for a long period of time (mean length of residence = 10.6 years, SD = 6.4, range: 5-27). Bilinguals self-rated their English proficiency on a 1-7 Likert scale and all indicated strong confidence with English (mean speaking = 5.8; mean listening = 6.2; mean reading = 6.3; mean writing = 5.8). Objective proficiency measures were obtained using a modified version of the Michigan ECPE, and all participants showed high proficiency in English (mean total score = 45.1 (50 max), SD = 4.2, range: 36-50).

Materials. Two hundred and forty sentence octuplets were created in a 2x2x2 design, crossing the factors of sentence grammaticality (grammatical, ungrammatical) based on verbal agreement with the head noun, attractor noun number (singular, plural), and modifier structure (PP, RC). PP and RC modifiers were matched in number of words, approximate semantic complexity and relatedness to the head noun. All sentences contained a singular head noun.

Some studies of syntactic processing have reported an additional left anterior negativity (LAN) preceding the P600 (e.g., Friederici, Hahne, & Mecklinger, 1996; Osterhout & Holcomb, 1992); however, the LAN is highly inconsistent across studies, even in native speakers, and is therefore not a reliable index of syntactic processing. Importantly, no LAN was found in the native speaker group in this study, and it therefore will not be considered further here (see Osterhout, McLaughlin, Kim, Greewald & Inoue, 2004 for discussion).
followed by a modifying phrase, followed by a verb (is/are, was/were, has/have), followed by a short predicate. Predicates were designed to be more plausible continuations of the head noun than attractor noun. An example octuplet is provided in (1), below. The eight versions of each sentence were distributed among eight experimental lists in a Latin square design, such that each list contained 30 sentences in each of the eight conditions. The experimental sentences were pseudo-randomized among 300 other sentences, half of which contained other types of syntactic or semantic anomalies. Overall, each list contained 540 sentences, half of which contained some sort of anomaly.

(1) a. The winner of the big trophy/trophies has/have proud parents.
    b. The winner who got the trophy/trophies has/have proud parents.

Procedure. 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 400ms followed by a 200ms blank screen. Sentence-ending words appeared with a full stop, followed by a response prompt asking for an acceptability judgment. Participants were asked to respond “yes” if they felt a sentence was well-formed and made sense in English or “no” if they felt the sentence was ungrammatical, implausible, or in any way unacceptable. “Yes” response hand was counterbalanced across subjects. Because of the large number of experimental sentences, participants took part in two ERP recording sessions. During a session, each participants saw half of the sentences from his or her respective list. ERPs were initially averaged within each session, but since there were no apparent differences across sessions, ERPs were re-averaged across both sessions for each participant. All data reported below reflect averages across both sessions.

Data acquisition and analysis. Continuous EEG was recorded from 19 scalp electrodes. Here we report only data from posterior midline electrode Pz, where P600 effects are usually largest (see Tanner, 2011 for full-scalp effects). 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 with a sampling rate of 250hz; 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 verbs italicized in (1), above), were averaged off-line for each subject at each electrode site, relative to a 100ms prestimulus baseline for native English speakers. Because of minor anomalies in the prestimulus baseline in the bilingual group, a mixed 50ms prestimulus to 50ms poststimulus baseline was used. P600 effects were quantified as mean amplitude difference between ungrammatical and grammatical conditions in two time windows: 500-700ms and 700-1000ms, corresponding to the early and late P600 effects, respectively (Hagoort & Brown, 2000). Amplitude differences were used as the dependent measure because of the large number of factors in this study, which could
potentially lead to a large number of statistical interactions. Since effects on P600 amplitude associated with attraction interference and structural modulation were the primary outcomes of interest, using amplitude differences allows us to see these effects while removing one layer of statistical interaction. Thus, the main effect of grammaticality (i.e., the presence of a P600 effect) was tested by the between-subjects intercept. Attractor number, and modifier structure were entered into an ANOVA as within-subjects factors and group (L1 English or Spanish-English bilingual) was a between-subjects factor. Because amplitude differences were used, all main effects should be interpreted as interactions with grammaticality (i.e., a modulation of the P600 effect elicited by ungrammatical verbs). Behavioral data were quantified as d-prime scores. A d-prime of 0 indicates chance performance distinguishing grammatical from ungrammatical sentences, while a d-prime of around 4 indicates near perfect discrimination. ANOVAs on d-prime scores included attractor number and modifier structure as within-subjects factors and group as a between-subjects factor.

3. Results

Behavioral results. D-prime scores and standard errors are presented for each condition and group in Table 1. Statistical analysis showed a main effect of group, such that monolingual English speakers were more sensitive to sentence grammaticality than bilinguals, $F(1, 49) = 13.539, p < .001$. There was a main effect of attractor number, $F(1, 49) = 34.772, p < .001$, indicating that sensitivity decreased when a plural attractor intervened between the head noun and verb (i.e., an attraction effect), and there was a main effect of modifier structure, $F(1, 49) = 8.333, p = .006$, indicating that sensitivity improved when the modifier was a RC, relative to a PP modifier. Importantly, there were no interactions between any of the factors (all Fs <1, except the structure by group interaction, $F(1, 49) = 1.434, p > .236$), indicating that the plural attractor effect was not modulated by modifier structure, and that both groups showed similar profiles of attraction interference and modulation by modifier structure.

<table>
<thead>
<tr>
<th>Attractor Number</th>
<th>English Monolingual</th>
<th>L1 Spanish-L2 English</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PP Modifier</td>
<td>RC Modifier</td>
</tr>
<tr>
<td>Singular</td>
<td>3.77 (.14)</td>
<td>3.96 (.16)</td>
</tr>
<tr>
<td>Plural</td>
<td>3.31 (.18)</td>
<td>3.41 (.18)</td>
</tr>
</tbody>
</table>

ERP Results. Grand mean waveforms for electrode Pz for both groups in each condition, comparing brain responses to grammatical (solid line) and ungrammatical (dashed line) verbs, are presented in Figure 1. Positive voltage is plotted down; vertical calibration bar represents $3\mu$V of activity. Plots show 100ms of prestimulus and 1000ms of poststimulus activity. Each tick mark represents 100ms of time.
Statistical analysis in the 500-700ms time window showed a main effect of grammaticality, indicated by a significant difference from zero in the intercept test, $F(1, 49) = 49.915, p < .001$, and a main effect of group, indicating that P600 effects were reliably smaller in the bilingual group than the monolingual group, $F(1, 49) = 8.101, p = .006$. There was a significant effect of attractor number, $F(1, 49) = 16.774, p < .001$, indicating that P600 effects were smaller following a plural versus singular attractor. No other effects or interactions were significant in this time window.

In the 700-1000ms time window, there was a main effect of grammaticality, indicated by a significant intercept test, $F(1, 49) = 87.295, p < .001$. There was an effect of group, $F(1, 49) = 9.125, p = .004$, again indicating that bilinguals’ P600 effects were significantly smaller than monolinguals’. There was also an effect of attractor number, $F(1, 49) = 26.425, p < .001$, signaling a reduction in P600 effect sizes following plural attractors. In this time window there was an additional effect of modifier structure, $F(1, 49) = 4.663, p = .036$, with P600 effects larger for ungrammatical verbs following RC than PP modifiers. Importantly, there were no further interactions between any of the factors ($Fs < 1.4$, $ps > .25$), suggesting that the structural effect was not modulated by attractor number and that the attraction and structural effect profiles were similar.
across both groups. Effects for each group and condition in the 700-1000ms time window are depicted in Figure 2. Positive deviations from zero indicate a P600 effect elicited by ungrammatical verbs.

**Figure 2. P600 effects (Pz) and std. errors between 700-1000ms**

4. Discussion

This experiment investigated the neural correlates of grammatical agreement processing in monolingual English speakers and late Spanish-English bilinguals, focusing in particular on the effects of attraction interference and structural complexity of the subject NP. Results showed that, relative to grammatical control verbs, disagreeing verbs elicited a late positive-going wave (i.e., a P600 effect). The overall size of P600 effects was smaller in native Spanish speakers, but in both groups it was smaller following plural attractor nouns (an attraction effect) and larger following RC modifiers than PP modifiers. Moreover, attractor number and modifier structure did not interact, suggesting that the structural and attraction effects are independent. These results show that previously reported reductions in ungrammaticality effects associated with agreement attraction found in behavioral experiments are also indexed in individuals’ brainwave activity. This study additionally shows that the grammatical complexity of the subject NP provides an additional modulation of sensitivity to agreement violations, both in behavioral (acceptability judgment) and electrophysiological measures. However, since the attraction and structural effects did not interact, we have no evidence that attraction is clause-bound, as has been shown to be the case in production. We return to a comparison of effects in the L1 and L2 groups below, but turn first to a discussion of theoretical accounts for the interference and structural effects found here.

While a full discussion of theoretical accounts of agreement attraction is beyond the scope of this paper, we will suggest that our results regarding
attraction interference and modulation of sensitivity to agreement violations by subject NP structure are most compatible with cue-based memory retrieval models of sentence comprehension (Van Dyke & McElree, 2011; Lewis & Vasishth, 2005; Lewis, Vasishth, & Van Dyke, 2006). A core component of these models is that they predict that during the establishment of a long-distance dependency, such as agreement, competition may arise when more than one candidate held in working memory resonates with the retrieval cues of the item triggering a search. With regard to attraction interference, this model predicts that when the verb disagrees with syntactic predictions established by the first NP (e.g., when a plural verb were is detected in the presence of a prediction for a singular verb following a singular head noun), a parallel, cue-based search in working memory occurs for possible controllers. A singular head noun may resonate with some retrieval cues triggered by the verb (e.g., [+nominative]), though a plural attractor noun will resonate with the [+plural] feature detected on the verb. Thus, the plural attractor noun will be incorrectly retrieved as the controller of agreement on a certain number of occasions, thereby reducing overall ungrammaticality effects (Tanner, 2011; Wagers et al., 2009; see also Badecker & Kuminiak, 2007).

Regarding structural effects, cue-based parsing predicts anti-locality effects, such that verbs are more easily integrated when more modificational material intervenes between a verb and an argument which must be integrated with that verb (Vasishth & Lewis, 2006). This follows from the model’s assumption that repeated reactivations of a constituent, associated with reprocessing during modification, raise that item’s base activation level. Higher base activation levels facilitate retrieval at an integration point. Evidence for anti-locality effects has come primarily from verb-final languages such as Hindi or German (Konieczny 2000; Vasishth & Lewis, 2006). In these studies, reading times for sentence-final verbs were shorter when the verbs were preceded by longer, more complex modifiers. The current study indicates that this effect may be present in head-initial languages, even when the length, but not structural complexity, of the modifier is identical across conditions. For structures used in the current experiment, the parser might establish a syntactic prediction for a singular verb upon encountering the first NP, and that prediction would become reinforced when processing the subject modifier. Processing a subject relative clause requires more retrievals of subject NP information and the associated syntactic predictions than the processing of a prepositional phrase because of verb-argument integration within the relative clause itself. Thus, the prediction of the matrix verb and its associated features is stronger when the verb has been preceded by a subject RC compared to a PP, and is therefore more easily integrated (see Tanner, 2011 for self-paced reading evidence in this regard).

In the current experiment, this has the effect of increasing participants’ sensitivity to an agreement violation because the expectation of the verb is stronger. This model therefore correctly predicts that the number feature of the attractor noun and structure of the subject modifier phrase will not interact when processing agreement in comprehension. Note, however, that structural effects
found in production involve an interaction between attractor number and subject NP structure, such that differences between PP and RC modifiers were only found when the attractor was plural; structural effects here were found both for singular and plural attractors. Additionally, reaction time evidence from production provided by Staub (2009) suggests that the interaction is not due to a ceiling effect from low error rates in the singular attractor conditions. The lack of interaction in this experiment suggests that effects of subject NP structure are therefore not influences on attraction itself, as has been shown to occur in production. Rather, there is a more general structural facilitation effect, similar to anti-locality effects shown in studies of head-final languages (Konieczny, 2000; Vasishth & Lewis, 2006; see Tanner, 2011 for a more extensive discussion of how cue-based parsing models relate to agreement attraction and the structural facilitation effect found here).

A final issue surrounds the effects of language group on the interference and structural facilitation effects found here. Recall that the SSH predicted that L2 learners should show very different profiles of attraction interference from native English speakers. If L2 learners adopt either linear or thematic strategies for grammatical agreement where NP1 is always the controller of agreement, they should show no effects of interference from plural attractors. Alternately, if learners always use a linear adjacency strategy in computing agreement, grammatical sentences with singular head and plural attractor nouns should show ungrammaticality effects because of the local disagreement between the attractor noun and verb. Moreover, it was predicted that if learners ignore grammatical structural information during sentence comprehension, there should be no difference in judgment accuracy or brain responses to verbs following PP versus RC modifiers. However, none of these predictions was borne out in the current data. While the late bilingual group showed overall smaller P600 effects than native speakers (cf. Rossi, Gugler, Friederici, & Hahne, 2006), the overall profiles of attraction interference and structural modulation were similar across the two groups. The current experiment thus allows two broad inferences about agreement processing in the group of late Spanish-English bilinguals tested here. First, the ERP correlates of agreement anomaly detection (i.e., P600 effects) were qualitatively similar across both the native and L2 English groups. Second, since P600 effects were modulated by the same factors and in the same ways, this suggests that similar information was consulted during the establishment of agreement dependencies in both the monolingual and L2 English speakers. The current results are thus in line with the proposal that, while there may be quantitative differences between L1 and L2 syntactic processing (e.g., smaller, later effects and more between-subjects variability), L1 and L2 processing may be qualitatively similar. Importantly, the data here indicate that late learners can go beyond strictly shallow linear- or thematically-based parsing strategies.
References


