Are L2 Learners Pressed for Time? Retrieval of Grammatical Gender Information in L2 Lexical Access

Kailen Shantz and Darren Tanner

1. Introduction

Even at high levels of proficiency, grammatical gender poses a persistent challenge to adult learners of a second language. While a considerable body of research has investigated L2 use and processing of grammatical gender, it remains unclear what the source of this difficulty is. Evidence does, however, suggest that it is not a knowledge deficit. Studies using offline comprehension tasks that probe knowledge of grammatical gender have, for example, repeatedly found that advanced L2 learners perform at ceiling (Alarcón, 2011; Grüter et al., 2012; McCarthy, 2008; Montrul et al., 2008; White et al., 2004). This has been taken as evidence that late L2 learners acquire the mental representations for grammatical gender information. Further evidence has come from research demonstrating online sensitivity to gender agreement. Self-paced reading and eye-tracking studies have shown that, like native speakers, late L2 learners spend more time reading a post-nominal adjective in Spanish that does not agree in gender with its preceding noun compared to adjectives that do agree (Foote, 2011; Keating, 2009; Sagarra & Herschensohn, 2010). Event-related potential (ERP) studies have further found that L2 learners can show native-like P600 responses to gender agreement violations (Alemán Bañon et al., 2014; Foucart & Frenck-Mestre, 2011, 2012; Gillon Dowens et al., 2010, 2011; Tokowicz & MacWhinney, 2005). Together, these findings that advanced adult L2 learners show ceiling performance in offline tasks as well as online sensitivity to gender agreement violations suggest that the source of their difficulty in using grammatical gender online does not exclusively stem from a knowledge deficit.

An alternative possibility is that L2 difficulty with grammatical gender has a processing-based locus. Indeed, there is substantial research to suggest that this is the case. Using eye-tracking while reading, Keating (2009) found that

---

* Kailen Shantz, University of Illinois at Urbana-Champaign, shantz2@illinois.edu. Darren Tanner, University of Illinois at Urbana-Champaign. We thank Nyssa Bulkes, Jack Dempsey, Lorenzo Grego, Chase Krebs, Yisi Liu, Maha Tolba, Rebekah Seyfert and Fiona Weingartner for help with data collection. This project was supported by a grant made to Darren Tanner from the Illinois Campus Research Board (RB14158). The first author received financial support from a Doctoral Fellowship from the Social Sciences and Humanities Research Council of Canada, and from an Illinois Distinguished Fellowship from the University of Illinois.

 Proceedings of the 40th annual Boston University Conference on Language Development, ed. Jennifer Scott and Deb Waughtal, 331-345. Somerville, MA: Cascadilla Press.

advanced late L2 learners of Spanish did not show sensitivity to gender agreement violations occurring outside of a determiner phrase, whereas these participants did show sensitivity within the determiner phrase; native speakers, in contrast, were sensitive to violations regardless of whether they occurred within the determiner phrase (cf. Gabriele et al., 2013). Studies have also shown variability in ERP responses elicited by gender agreement violations for late L2 learners compared to native speakers. While native speakers consistently show P600 responses to gender agreement violations in sentential contexts (Barber & Carreiras, 2005; Davidson & Indefrey, 2009; Demestre & Garcia-Albea, 2007; Gunter et al., 2000 Hagoort, 2003; Hagoort & Brown, 1999; Molinaro et al., 2008; Nevins et al., 2007; O'Rourke & Van Petten, 2011; Sabourin & Haverkort, 2003; Wicha et al., 2004), some studies fail to find any P600 response to adjective-noun agreement violations in late L2 learners (Foucart & Frenck-Mestre, 2011, 2012; Morgan-Short et al., 2010).

In addition to variable sensitivity to gender agreement, late L2 learners also show a reduced or absent ability to use gender marking information in order to anticipate upcoming nouns in an utterance. For example, Guillelmon and Grosjean (2001) had native speakers as well as early and late L2 speakers of French perform a repetition task in which they listened to auditorily presented noun phrases and were asked to repeat the nouns they heard. The determiner that preceded the noun in each noun phrase was either congruent in gender with the noun, incongruent in gender, or gender-neutral. For native speakers and early L2 acquirers of French, naming latencies were shorter when a noun was preceded by a congruent determiner than when preceded by a gender-neutral determiner, suggesting that gender information on the determiner facilitated lexical access to the nouns. For these same groups, naming latencies were longer when nouns were preceded by an incongruent determiner than when preceded by a neutral determiner, indicating that these participants may have been using gender information on the determiner to anticipate the upcoming noun, and experienced a processing cost when their expectations were not met. In contrast to these results, late L2 learners of French showed no effect of gender congruency on their naming latencies, suggesting that they were unable to or inefficient at using gender information on the determiners to anticipate the upcoming nouns.

A similar discrepancy between L1 and late L2 processing has been found in looking-while-listening and eye-tracking research. In these tasks, participants hear phrases containing a gender-marked determiner and a target noun while they view a visual display of nouns belonging either to different gender classes or to the same gender class. Native speakers are consistently faster at orienting to a target image on trials where gender marking is an informative cue to the identity of a target noun compared to when it is an uninformative cue (Dussias et al., 2013; Grüter et al., 2012; Hopp, 2013, 2015; Lew-Williams & Fernald, 2007, 2010). L2 learners, in contrast, are often no faster at orienting to the target noun on informative trials than on uninformative trials (Grüter et al., 2012; Lew-Williams & Fernald, 2010), though Hopp (2013, 2015) found that advanced L2 learners of German who were more accurate in a gender assignment task did show anticipatory looks on informative trials. Hopp (2013) further reports evidence that individuals' speed of lexical access predicts whether they will exhibit anticipatory looks (see also Dussias et al., 2013).

An aspect of L2 processing that has, to our knowledge, been unexplored is the time course of lexical access. Prior research on L1 populations has found that grammatical gender is retrieved prior to phonology (Van Turennout et al., 1998). Additional research has shown, however that the time course of information retrieval during lexical access is sensitive to retrieval difficulty (Abdel Rahman et al., 2003; Abdel Rahman & Sommer, 2003). Specifically, these studies showed that the timing of semantic information retrieval relative to phonological retrieval is delayed when semantic retrieval is made more difficult, without affecting the timing of phonological retrieval. It is therefore plausible that L2 speakers may experience delayed retrieval of grammatical gender information relative to phonological information, if it is more difficult for them to retrieve grammatical gender. Such delayed retrieval might, in turn, explain some of the difficulties L2 populations experience in using grammatical gender online to anticipate or compute agreement relations. The current study therefore employs a covert naming task in order to investigate whether late L2 learners of German are delayed in their retrieval of grammatical gender information relative to their retrieval of phonological information.

2. Current Study

2.1. Method

To investigate the relative time course of access to phonology and gender in L2 lexical retrieval, we combined the use of EEG with a dual-choice go/no-go task. We followed previous research using this paradigm (e.g. Van Turennout et al., 1998; Schmitt et al., 2000) in focusing on two ERP components as measures for estimating the relative time course of lexical access: the lateralized readiness potential (LRP) and the N200. The LRP is an ERP effect that, in dual-choice contexts with a left/right hand response, reflects preparation to respond with a specific hand. The LRP begins to develop as soon as task-relevant information becomes available to be passed on to the motor system (Kutas & Donchin, 1980). Importantly, in go/no-go tasks the LRP may also develop briefly on no-go trials before a motor response is inhibited by the decision-making system. Relative time course information is obtained based on whether an LRP develops on no-go trials when one source of information determines the hand response, but not the other. The rationale is that if one source of information (e.g. gender) is retrieved before the other (e.g. phonology), an LRP will develop on no-go trials when that earlier source of information (i.e. gender) determines the hand response. The no-go LRP is then later aborted when the go/no-go information source is retrieved and can inform the decision making system that no response is needed. The LRP thus provides an upper estimate of when the linguistic feature determining response hand was retrieved.

The other ERP component under investigation, the N200, is a negative-going potential over frontal sites that is larger on no-go trials than on go trials and is assumed to reflect response inhibition (Folstein & Van Petten, 2008). As such, the N200 complements the LRP by providing an upper estimate of when the linguistic feature determining the go/no-go response was retrieved. Relative time course information is inferred by comparing N200 latencies across conditions where the go/no-go response is determined by different sources of information (e.g. gender or phonology).

In addition to the dual-choice go/no-go task, we also included a gender decision task to assess individuals' speed of access to grammatical gender information in the absence of a secondary task.

2.2. Participants

Nineteen L1-English L2 learners of German with normal or corrected-to-normal vision participated in this experiment (10 female, 8 left-handed)[1]. Table 1 summarizes L2 participants' demographic and proficiency data. Participants also included twenty native speakers of German with normal or corrected-to-normal vision (7 female, 3 left-handed, age 20-35, *M* = 24.5). No participants reported any history of neurological impairments, and all provided informed consent and received modest monetary compensation for their participation.

Table 1. Summary of L2 participants' demographic and German proficiency data (n = 19).

	Mean	Std. Dev	Range
Age	21.4	3.1	18-30
AoA	13.8	1.0	11-16
Proficiency Measure			
Goethe Institut Test (30 max)	17.9	4.1	12-25
Self-Rated German Proficiency			
Speaking (10 max)	6.7	1.5	4-9
Listening (10 max)	7.2	1.8	4-10
Writing (10 max)	6.6	1.2	4-8
Reading (10 max)	7.1	1.4	5-9

2.3. Materials
2.3.1. Dual-Choice Go/No-Go Task

Materials for the dual-choice go/no-go task were 24 black-and-white images taken from the International Picture Naming Database (Bates et al., 2000). All images depicted high frequency, concrete nouns with high naming agreement in German. Nouns corresponding to the images were equally balanced across two grammatical gender categories (masculine and neuter) and two word-initial

[1] Data from one additional participant were removed due to a recording issue.

phone categories (/b/-initial and /k/-initial). Thus there were six masculine /b/-initial items, six masculine /k/-initial items, six neuter /b/-initial items and six masculine /k/-initial items. No nouns had natural gender or any morphological, orthographic or phonological markers of gender class.

2.3.2. Gender Decision Task

Items in the gender decision task consisted of 60 high frequency and 60 mid frequency nouns that were morphologically simple and did not have natural gender. Within each frequency band there were 20 nouns with masculine gender, 20 with feminine gender and 20 with neuter gender. Each of these subsets of 20 nouns were matched for character length.

2.4. Procedure

All experimental tasks were administered using Paradigm (Perception Research Systems, 2007). Data collection took place in a single session lasting between 180 and 210 minutes. Each session began with a language history questionnaire and an abridged version of the Edinburgh Handedness Inventory. This was followed by a brief warm-up in which the images used in the dual-choice go/no-go task were randomly presented on a computer screen to participants, who were asked to provide the name and grammatical gender of each item. This was done to ensure that participants knew the gender for each item and were supplying the expected names for each image.

After the warm-up, participants completed the dual-choice go/no-go task. In this task, the images were presented to participants individually in a central viewing position on a computer monitor. Participants were asked to use one source of information about the depicted noun (either the gender or initial phone) to decide whether or not to press a button (i.e. the go/no-go portion of the task), and the other source of information to decide whether to respond with their right hand or left hand on go trials. We manipulated which information source determined which decision within subjects (i.e. whether grammatical gender or phonology determined the go/no-go response or the hand response). Participants encountered each possible configuration in separate blocks. Response hand was also counterbalanced within subjects. Thus each participant encountered eight blocks total. Within a block, items were presented in a randomized order over four repetitions for a total of 96 trials per block. An item was not repeated until all 24 items had been presented in a single repetition of item presentation. Block order was counterbalanced across four lists, and list assignment was balanced across participants so that an equal number of participants saw each list.

Each trial began with a fixation cross lasting 300 ms and a blank screen lasting 200 ms, followed by an image in the center of the computer screen. Images remained on screen for a maximum of two seconds, during which time participants made their go/no-go responses on a handheld gamepad using their left and right index fingers. Trials terminated when a response was made or after the two second timeout period. Upon termination of a trial, there was an 800 ms blank screen followed by a screen lasting 1700 ms that participants had been instructed to use for blinking, then a 300 ms blank screen; then the next trial began. Participants received short breaks between blocks, as well as instructions for the next block. Each block began with a practice in which participants responded to each experimental item following the new block instruction. Participants repeated the practice as many times as necessary until they were confident with the new configuration.

Following the dual-choice go/no-go task, participants completed the gender decision task. Items in this task were presented in a randomized order, and participants responded by pressing one of three pre-specified buttons on the gamepad to indicate if the gender of a word was feminine, masculine or neuter. Each trial began with a fixation cross in the center of the screen lasting for 300 ms, followed by a blank screen of 200 ms and then the string of letters, which remained on screen until a response was made. After participants made their responses, there was a blank screen with a variable interstimulus interval between 1500 and 1800 ms before the next trial began.

2.5. Data Acquisition and Analysis

Continuous EEG was recorded from 28 tin scalp electrodes mounted in an elastic cap (Electro-cap International), in accordance with the extended 10-20 system (Jasper, 1958). Eye movements and blinks were monitored with electrodes placed below the left eye and at the outer canthus of each eye. The EEG was amplified with a BrainAmpDC bioamplifier system (Brain Products, Gilchin, Germany) and digitized with a 1000-Hz sampling rate and an online 450 Hz lowpass filter and a 0.016 Hz high pass filter (10s time constant). Impedances were held below 5 kΩ at scalp and mastoid sites and below 15 kΩ at the ocular electrodes. All electrodes were referenced online to the left mastoid and re-referenced offline to the algebraic mean of activity over the left and right mastoids.

Offline processing was conducted with EEGLAB (Delorme & Makeig, 2004) and ERPLAB (Lopez-Calderon & Luck, 2014) toolboxes. Only trials with correct responses were processed for analysis. A 0.1- to 30-Hz bandpass zero phase-shift Butterworth IIR filter (-6-dB cutoff, 12-dB/octave roll-off) was applied offline to the continuous EEG. Trials with excessive artefacts not related to eye movements and blinks, or with ocular activity earlier than 250 ms post-stimulus were hand-rejected. Independent component analysis (ICA) was then used to isolate and remove EEG components related to blinks and eye movements on remaining trials. After ICA artefact correction was completed, trials were rescreened to remove epochs with excessive drift, muscle activity or other artefacts. In total, 5.1% of trials were rejected for the L2 participants, and 4.2% of trials for the L1 participants. Data were then downsampled to 250 Hz[2].

ERPs were calculated over a 1800 ms window time-locked to the onset of the picture for each participant over each electrode for each condition, relative to a 200 ms pre-stimulus baseline. LRPs were calculated by subtracting the averaged EEG that was ipsilateral to a response hand from the averaged contralateral activity to a response hand at electrode sites C3 and C4, which are located near the primary motor cortices (see Kutas & Donchin, 1974). LRPs were quantified by averaging 4 ms time bins from 0-1000 ms. Bins were averaged separately for each combination of condition (go/no-go = phonology and hand = gender or go/no-go = gender and hand = phonology), trial type (go or no-go) and participant, yielding four data points per participant per time bin. N200s were quantified as peak latency in the no-go minus go difference waves at Fz in the 200-700 ms time window for each combination of condition, trial type and participant. N200 peak latencies were determined using a jackknife procedure (Miller et al., 1998).

All statistical analyses were carried out in R (R Core Team, 2015). N200 data were analyzed with ANOVAs using the *F*-value correction procedure for jackknifing following Ulrich & Miller (2001). LRPs were analyzed as a time series by fitting the averaged time bins to a generalized additive mixed model (GAMM; see Wood, 2006). GAMMs are a semi-parametric extension of linear regression that allow the modeler to fit smooths to the data in order to model non-linearity. GAMMS provide *p*-values for smooth terms as an indication of whether the smooths are significantly different from zero. Interpretation of effects, however, requires visualization.

Behavioral data were analyzed with mixed effects models using linear models for the reaction time data and generalized linear models for the accuracy data. Only correct responses were analyzed for the reaction time data. Responses in the gender decision task were removed if they were shorter than 200 ms, longer than 3000 ms, or greater than 2.5 standard deviations from a participant's mean response time. This removed 0.8% of correct trials from the L1 data and 3.5% from the L2 data. No trials required trimming from the dual-choice go/no-go task. All models were fit with the maximal random effects structure that would converge on the data (Barr et al., 2013). Effects in the linear models are considered significant at an absolute *t*-value of 2 (Baayen et al., 2008).

3. Results

3.1. Behavioral Results

In the gender decision task, mean accuracy was 96.8% (*SD* = 2.4%) for the native speakers and 63.2% (*SD* = 12.7%) for the L2 speakers. Modeling results found this difference to be significant (*logOdds* = -3.00, $p < 0.001$). The mean

[2] EEG for one L2 participant was recorded at 250 Hz and was therefore not downsampled.

reaction time for the native speakers was 981.9 ms (*SD* = 154.1 ms). For the L2 participants this was 1283.0 ms (*SD* = 181.3 ms). This difference also proved significant (*ß* = 0.27, *t* = 5.83). The L1 speakers of German were thus significantly faster and more accurate at assigning grammatical gender to German nouns than the L2 participants.

Mean accuracy for the native speakers in the dual-choice go/no-go task was 97.8% (*SD* = 1.7%) on go trials and 99.3% (*SD* = 0.7%) on no-go trials. L2 participants were 98.3% (*SD* =1.4%) accurate on go trials and 99.6% (*SD* = 0.6%) accurate on no-go trials. No effects of group, i.e. L1 or L2 speaker (*logOdds* = 0.31, *p* = 0.204), or condition, i.e. hand = phonology or hand = gender (*logOdds* = 0.12, *p* =0.321) emerged from the model fit to this data. Participants were thus highly accurate overall, regardless of condition or whether they were L1 or L2 speakers.

Mean response times for the native speakers were 942.2 ms (*SD* = 160.2) when hand = gender, and 946.2 ms (*SD* = 183.2 ms) when hand = phonology. L2 speakers' mean response times were 895.8 ms (*SD* = 156.3 ms) when hand = gender, and 930.0 ms (*SD* = 148.2 ms) when hand = phonology. The model for this data revealed no effect of group (*ß* = -0.02, *t* =-0.42) or condition (*ß* = 0.02, *t* = 0.84). In short, L2 and L1 participants were equally accurate and fast in the dual-choice go/no-go task, regardless of condition.

3.2. EEG Results

3.2.1. N200

Difference waves for the grand-average N200s are shown in Figure 1 in each condition for L1 and L2 participants. Mean peak latencies for the native speakers were 398.8 ms when go/no-go = gender and 419.0 ms when go/no-go = phonology. These latency differences did not emerge as significant in an ANOVA test (F (1,19) = 13.83, $F_{\text{CORRECTED}}$ = 0.04, $p_{\text{CORRECTED}}$ = 0.847). Mean peak latencies for the non-native speakers were 422.1 ms when go/no-go = gender and 419.4 ms when go/no-go = phonology. A type III ANOVA revealed no significant differences between these latencies (F (1,18) = 9.94, $F_{\text{CORRECTED}}$ = 0.03, $p_{\text{CORRECTED}}$ = 0.863).

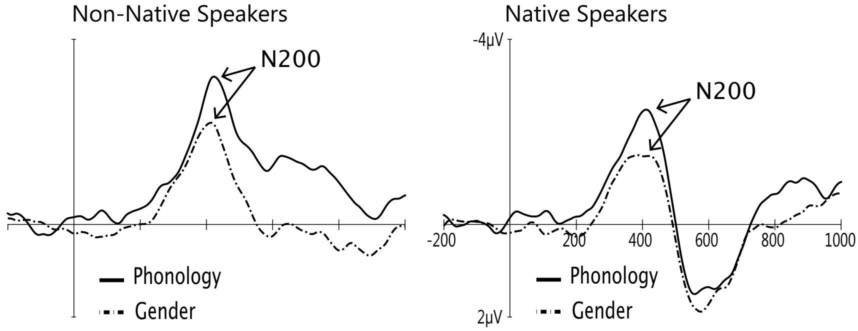

Figure 1. ERP plots of N200 effects for native and non-native speakers when go/no-go = phonology and go/no-go = gender. 200 ms of prestimulus and 1000 ms of poststimulus activity are depicted; negative voltage is plotted up.

A type III two-way ANOVA with the factors group (L1 or L2) and condition (go/no-go = gender or go/no-go = phonology) further yielded no significant main effect for group (F (1,37) = 13.18, $F_{\text{CORRECTED}} = 0.04$, $p_{\text{CORRECTED}} = 0.850$), nor did it find a significant interaction between group and condition (F (1,37) = 20.88, $F_{\text{CORRECTED}} = 0.06$, $p_{\text{CORRECTED}} = 0.811$)[3]. Figure 2 overplots the N200 difference waves for L1 and L2 participants by condition.

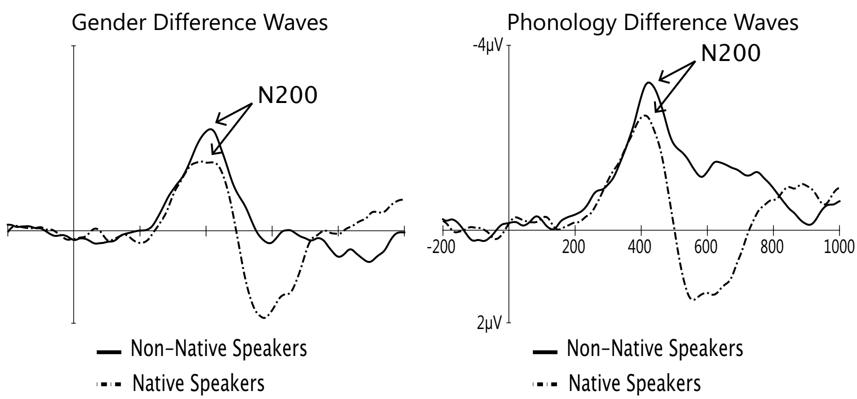

Figure 2. Overplot of difference waves when go/no-go = gender and when go/no-go = phonology for native and non-native speakers. 200 ms of prestimulus and 1000 ms of poststimulus activity are depicted; negative voltage is plotted up.

---

[3] Note that the F-correction procedure requires dividing the F-statistic by the square of the number of participants in a group minus one. Due to the unbalanced number of participants in each group, it was unclear what number to select for this. We chose to report the more conservative p-values obtained by dividing the F-statistics by $19^{\wedge 2}$.

3.2.2. LRP

Figure 3 shows the grand-average LRPs for the L1 and L2 participants in each condition. GAMM results find no significant no-go activity for L1 participants (Hand = Gender: *edf* = 1.0, $F = 1.81$, $p = 0.179$; Hand = Phonology: *edf* = 1.0, $F = 0.36$, $p = 0.547$) nor for the L2 participants (Hand = Gender: *edf* = 1.0, $F = 2.05$, $p = 0.152$; Hand = Phonology: *edf* = 1.0, $F = 0.28$, $p = 0.600$). The GAMM further revealed that go LRPs when gender = hand onset significantly earlier for the L2 participants (~368) ms compared to the native speakers (~428 ms). Figure 4 overplots these effects. The LRP data thus find no evidence that late L2 learners are delayed in their retrieval of grammatical gender relative to phonology or relative to native speakers.

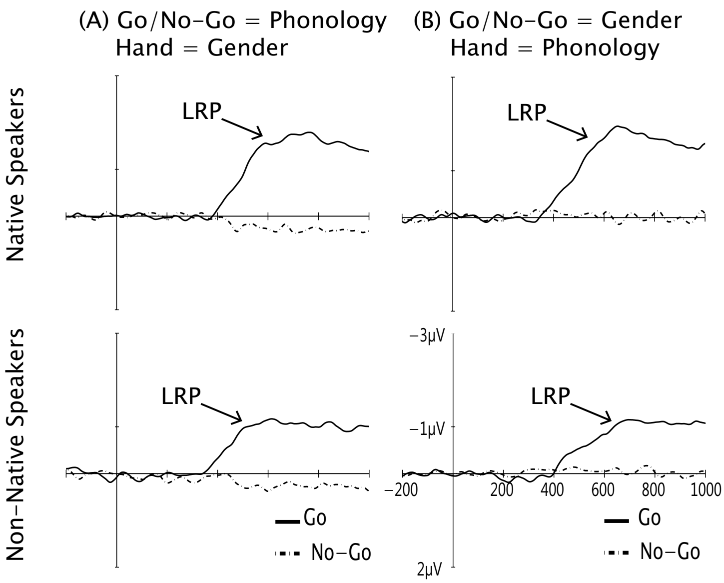

Figure 3. ERP plots of LRP effects for native and non-native speakers. Panel A shows LRPs when gender determines the hand response and panel B when phonology determines this response. 200 ms of prestimulus and 1000 ms of poststimulus activity are depicted; negative voltage is plotted up.

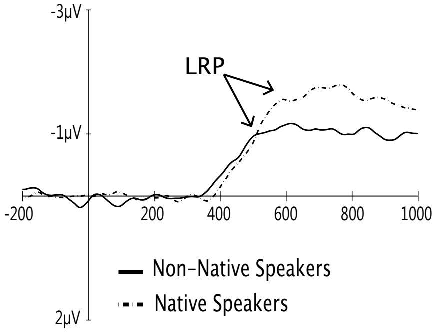

Figure 4. Overplot of go LRPs for native and non-native speakers when hand= gender. 200 ms of prestimulus and 1000 ms of poststimulus activity are depicted; negative voltage is plotted up.

4. Discussion

This study examined whether L2 processing difficulties with grammatical gender may stem partly from delayed retrieval of gender information during lexical access. Using a dual-choice go/no-go task, we found that late L2 learners of German are neither delayed in accessing grammatical gender relative to native speakers nor relative to phonology. We further found no evidence for earlier access to gender compared to phonology in our native speakers, in contrast to the findings of Van Turennout et al. (1998). Our results thus suggest that the difficulties L2 learners experience with grammatical gender cannot be directly tied to delays in the timing of gender access during online processing.

The ERP and behavioral data revealed a similar time course for access to gender and phonology for L1 and L2 speakers of German. N200 data showed no difference in peak latencies for grammatical gender or phonology when comparing L1 to L2 participants. Both L1 and L2 groups further showed an absence of no-go LRPs in both conditions. While this absence of no-go LRPs makes it difficult to draw any conclusions from a between-group comparison, the L2 group did show a significantly earlier go LRP when hand = gender compared to the L1 group. This suggests that the L2 participants were not delayed relative to the L1 participants in retrieving grammatical gender. The behavioral data provide additional support for this conclusion in revealing no differences in response times or accuracy between L1 and L2 speakers. Our data thus indicate that for our L1-English L2-German speakers, the lack of grammatical gender in the L1 does not preclude rapid access – on a time-scale similar to native Germans' – to gender information, at least for familiar nouns.

Critically, despite comparable performance in the dual-choice go/no-go task, L2 speakers were reliably less accurate and slower than native speakers in selecting the appropriate grammatical gender for a word, as found by the gender decision task. Thus the fact that L2 performance did not differ statistically from

L1 performance in the dual-choice go/no-go task cannot be attributable to high L2 proficiency or to L1 attrition.

In addition to finding a similar time course for access to grammatical gender and phonology across L1 and L2 participants, results further found no evidence for earlier access to grammatical gender over phonology in our L1 or L2 speakers of German. The N200 data revealed no differences in peak latencies between grammatical gender and phonology for either group, and there was a complete absence of no-go LRPs in both conditions for L1 and L2 participants. These results fail to replicate the findings of Van Turennout et al. (1998) that gender is retrieved before phonology in L1 lexical access. Instead, our results indicate that retrieval of grammatical gender and phonology follow a similar time course in both L1 and L2 lexical access.

In short, the current study found that access to grammatical gender and phonological information show a similar time course in both L1 and L2 lexical access. This finding places constraints on psycholinguistic accounts of the dissociation between online and offline L2 gender use, suggesting that this difficulty is not due to delayed access for familiar nouns. We further found no evidence for earlier access to grammatical gender over phonology, indicating that earlier access to grammatical gender may not be a pervasive feature of the language production system for native speakers, in contrast to earlier findings.

References

Abdel Rahman, Rasha & Sommer, Werner. (2003). Does phonological encoding in speech production always follow the retrieval of semantic knowledge? Electrophysiological evidence for parallel processing. *Cognitive Brain Research, 16,* 372-382.

Abdel Rahman, Rasha, Van Turennout, Miranda, & Levelt, Willem J. (2003). Phonological encoding is not contingent on semantic feature retrieval: An electrophysiological study on object naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 850-860.

Alarcón, Irma V. (2011). Spanish gender agreement under complete and incomplete acquisition: Early and late bilinguals' linguistic behavior within the noun phrase. *Bilingualism: Language and Cognition, 14*, 332-350.

Alemán Bañon, José, Fiorentino, Robert, & Gabriele, Alison. (2014). Morphosyntactic processing in advanced second language (L2) learners: An event-related potential investigation of the effects of L1-L2 similarity and structural distance. *Second Language Research, 30*, 275-306.

Baayen, R. H., Davidson, D. J., Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390-412.

Barber, Horacio & Carreiras, Manuel. (2005). Grammatical gender and number agreement in Spanish: An ERP comparison. *Journal of Cognitive Neuroscience, 17*, 137-153.

Barr, Dale J., Levy, Roger, Scheepers, Christoph, & Tily, Harry J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*, 255-278.

Bates, Elizabeth, Federmeier, Kara, Herron, Dan, Iyer, Gowri, Jacobsen, Thomas, Pechman, Thomas, … & Pléh, Csaba. (2000). Introducing the CRL International Picture-Naming Project (CRL-IPNP). *Center for Research in Language Newsletter, 12(1)*. Retrieved from http://crl.ucsd.edu/bates/papers/pdf/from-meiti/11-Picnorm.pdf.

Davidson, Douglas J. & Indefrey, Peter. (2009). An event-related potential study on changes of violation and error responses during morphosyntactic learning. *Journal of Cognitive Neuroscience, 21*, 433-446.

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

Demestre, Josep & García-Albea, José E. (2007). ERP evidence for the rapid assignment of an (appropriate) antecedent to PRO. *Cognitive Science, 31*, 342-354.

Dussias, Paola E., Valdés Kroff, Jorge R., Guzzardo Tamargo, Rosa E., & Gerfen, Chip. (2013). When gender and looking go hand in hand. *Studies in Second Language Acquisition, 35*, 353-387.

Folstein, Jonathan R. & Van Petten, Cyma. (2008). Influence of cognitive control and mismatch on the N2 component of the ERP: A review. *Psychophysiology, 45*, 152-170.

Foote, Rebecca. (2011). Integrated knowledge of agreement in early and late English-Spanish bilinguals. *Applied Psycholinguistics, 32*, 187-220.

Foucart, Alice & Frenck-Mestre, Cheryl. (2011). Grammatical gender processing in L2: Electrophysiological evidence of the effect of L1-L2 syntactic similarity. *Bilingualism: Language and Cognition, 14*, 379-399.

Foucart, Alice & Frenck-Mestre, Cheryl. (2012). Can late L2 learners acquire new grammatical features? Evidence from ERPs and eye-tracking. *Journal of Memory and Language, 66*, 226-248.

Gabriele, Alison, Fiorentino, Robert, & Alemán Bañon, José. (2013). Examining second language development using event-related potentials: A cross-sectional study on the processing of gender and number agreement. *Linguistic Approaches to Bilingualism, 3*, 213-232.

Gillon Dowens, Margaret, Guo, Taomei, Guo, Jingjing, Barber, Horacio, & Carreiras, Manuel. (2011). Gender and number processing in Chinese learners of Spanish – Evidence from event related potentials. *Neuropsychologia, 49*, 1651-1659.

Gillon Dowens, Margaret, Vergara, Marta, Barber, Horacio A., & Carreiras, Manuel. (2010). Morphosyntactic processing in late second-language learners. *Journal of Cognitive Neuroscience, 22*, 1870-1887.

Goethe Institut. (2004). *German Placement Test*. Retrieved from http://www.goethe.de/cgi-bin/einstufungstest/einstufungstest.pl (October, 2006).

Grüter, Theres, Lew-Williams, Casey, & Fernald, Anne. (2012). Grammatical gender in L2: A production or a real-time processing problem? *Second Language Research, 28*, 191-215.

Guillelmon, Delphine & Grosjean, François. (2001). The gender marking effect in spoken word recognition: The case of bilinguals. *Memory & Cognition, 29*, 503-511.

Gunter, Thomas C., Friederici, Angela D., & Schriefers, Herbert. (2000). Syntactic gender and semantic expectancy: ERPs reveal early autonomy and late interaction. *Journal of Cognitive Neuroscience, 12*, 556-568.

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

Hagoort, Peter & Brown, Colin M. (1999). Gender electrified: ERP evidence on the syntactic nature of gender processing. *Journal of Psycholinguistic Research, 28*, 715-728.

Hopp, Holger. (2013). Grammatical gender in adult L2 acquisition: Relations between lexical and syntactic variability. *Second Language Research, 29*, 33-56.

Hopp, Holger. (2015). *Learning (not) to predict: Grammatical gender processing in adult L2 acquisition.* Manuscript submitted for publication.

Jasper, Herbert H. (1958). Report of the Committee on Methods of Clinical Examination in Electroencephalography: 1957. *Electroencephalography and Clinical Neurophysiology, 10*, 370-375.

Keating, Gregory D. (2009). Sensitivity to violations of gender agreement in native and nonnative Spanish: An eye-movement investigation. *Language Learning, 59*, 503-535.

Kutas, Marta & Donchin, Emanuel. (1974). Studies of squeezing: Handedness, responding hand, response force, and asymmetry of readiness potential. *Science, 186*, 545-548.

Kutas, Marta & Donchin, Emanuel. (1980). Preparation to respond as manifested by movement-related brain potentials. *Brain Research, 202*, 95-115.

Lew-Williams, Casey & Fernald, Anne. (2007). Young children learning Spanish make rapid use of grammatical gender in spoken word recognition. *Psychological Science, 18*, 193-198.

Lew-Williams, Casey & Fernald, Anne. (2010). Real-time processing of gender-marked articles by native and non-native Spanish speakers. *Journal of Memory and Language, 63*, 447-464.

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

McCarthy, Corrine. (2008). Morphological variability in the comprehension of agreement: An argument for representation over computation. *Second Language Research, 24*, 459-486.

Miller, Jeff, Patterson, Tui, & Ulrich, Rolf. (1998). Jackknife-based method for measuring LRP onset latency differences. *Psychophysiology, 35*, 99-115.

Molinaro, Nicola, Vespignani, Francesco, & Job, Remo. (2008). A deeper reanalysis of a superficial feature: An ERP study on agreement violations. *Brain Research, 1228*, 161-176.

Montrul, Silvina, Foote, Rebecca, & Perpiñán, Silvia. (2008). Gender agreement in adult second language learners and Spanish heritage speakers: The effect of age and context of acquisition. *Language Learning, 58*, 503-553.

Morgan-Short, Kara, Sanz, Cristina, Steinhauer, Karsten, & Ullman, Michael T. (2010). Second language acquisition of gender agreement in explicit and implicit training conditions: An event-related potential study. *Language Learning, 60*, 154-193.

Nevins, Andrew, Dillon, Brian, Malhotra, Shiti, & Phillips, Colin. (2007). The role of feature-number and feature-type in processing Hindi verb agreement violations. *Brain Research, 1164*, 81-94.

O'Rourke, Polly L. & Van Petten, Cyma. (2011). Morphological agreement at a distance: Dissociation between early and late components of the event-related brain potential. *Brain Research, 1392*, 62-79.

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

Perception Research Systems. (2007). Paradigm Stimulus Presentation (Version 2.3.0.29) [computer software]. Retrieved from http://www.paradigmexperiments.com.

R Core Team. (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing (Version 2.3.3) [computer software]. Vienna, Austria. Retrieved from http://www.R-project.org/.

Sabourin, Laura & Haverkort, Marco. (2003). Neural substrates of representation and processing of a second language. In Roeland van Hout, Aafke Hulk, Folkert Kuiken, & Richard Towell (Eds.), *The Lexicon-Syntax Interface in Second Language Acquisition* (pp. 175-195). Amsterdam: Benjamins.

Sagarra, Nuria & Herschensohn, Julia. (2010). The role of proficiency and working memory in gender and number agreement processing in L1 and L2 Spanish. *Lingua, 120*, 2022-2039.

Schmitt, Bernadette M., Münte, Thomas F., & Kutas, Marta. (2000). Electrophysiological estimates of the time course of semantic and phonological encoding during implicit picture naming. *Psychophysiology, 37*, 473-484.

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*, 173-204.

Ulrich, Rolf & Miller, Jeff. (2001). Using the jackknife-based scoring method for measuring LRP onset effects in factorial designs. *Psychophysiology, 38*, 816-827.

Van Turennout, Miranda, Hagoort, Peter, & Brown, Colin M. (1998). Brain activity during speaking: From syntax to phonology in 40 milliseconds. *Science, 280*, 572-574.

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

Wicha, Nicole Y. Y., Moreno, Eva M., & Kutas, Marta. (2004). Anticipating words and their gender: An event-related brain potential study of semantic integration, gender expectancy, and gender agreement in Spanish sentence reading. *Journal of Cognitive Neuroscience, 16*, 1272-1288.

Wood, Simon N. (2006). *Generalized Additive Models. An Introduction with R*. Boca Raton, FL: Chapmans and Hall/CRC.

Proceedings of the 40th annual Boston University Conference on Language Development

edited by Jennifer Scott
and Deb Waughtal

Cascadilla Press Somerville, MA 2016

Copyright information

ISSN 1080-692X
ISBN 978-1-57473-066-1 (2 volume set, paperback)
ISBN 978-1-57473-166-8 (2 volume set, library binding)

Ordering information

To order a copy of the proceedings or to place a standing order, contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, sales@cascadilla.com, www.cascadilla.com