

Suprasegmental information cues morphological anticipation
during L1/L2 lexical access

Abstract

We use visual-world eye-tracking and gating methods to investigate whether Spanish monolinguals and English late learners of Spanish use prosodic cues (lexical stress) to anticipate morphological information (suffixes) during spoken word recognition, and if they do, whether L2 proficiency and working memory (WM) mediate their anticipatory abilities. Our findings show that the monolinguals used prosodic information to predict word endings in both tasks, regardless of first-syllable stress (stressed, unstressed) and structure (CV, CVC). In contrast, the beginning learners do not use prosodic information to anticipate word suffixes in any tasks or conditions. Importantly, the advanced learners mirrored the monolinguals, except in words with first-syllable CV structure, but were slower than the monolinguals. Finally, WM was not associated with anticipatory eye movements, though results were inconclusive for offline processing. Taken together, the present study shows that suprasegmental information facilitates morphological anticipation during spoken word recognition, and that adult learners can adapt native-like anticipatory processing strategies.

Keywords: morphological processing, anticipation, prediction, L2 proficiency, working memory, critical period, prosody, lexical stress

**Suprasegmental information cues morphological anticipation
during L1 and L2 lexical access**

Anticipation, the pre-activation of upcoming information, is considered to be an integral part of human information processing that allows comprehension to occur quickly and incrementally (Altman & Mirkovic, 2009; Dell & Chang, 2014; Pickering & Garrod, 2013). The importance of anticipation during sentence processing in monolinguals has been known for decades (see Kamide, 2008; Huettig, Rommers, & Meyer, 2011, for overviews). In contrast, L2 anticipatory studies have emerged in the last 5 years. L2 scholars agree that intermediate learners make reduced or no use of morphosyntactic cues to pre-activate a following referent in reading (Foucart, Martin, Moreno, & Costa, 2014; Grüter, Lew-Williams, & Fernald, 2012) and listening tasks (Foucart, Ruiz-Tada, & Costa, 2016; Hopp, 2015). However, they diverge in whether advanced learners can use morphosyntactic cues for anticipatory processes like natives do (accessibility accounts) (Marull, 2017), or not (deficit accounts) (Hopp, 2015; Martin et al., 2013).

Importantly, these L2 anticipatory studies have focused on morphological cues to predict upcoming words. Prosody is essential to process sentences (Cutler, Dahan, & van Donselaar, 1997) and words (McQueen, 2005). However, most L1 anticipatory studies on prosodic cues concentrate on the prediction of upcoming words and structures (see Hawthorne & Gerken, 2014, for a review). And the few studies on prosodic cues to anticipate morphology within a word are limited to two cues: Swedish word tones in monolinguals (Roll, Horne, & Lindgren, 2010; Roll, 2015; Söderström, Horne, & Roll, 2017; Söderström, Roll, & Horne, 2012), and

English vowel duration in monolinguals and adult learners (Rehrig, 2017). We extend this research to Spanish lexical stress cues in monolinguals and late learners.

Specifically, we use visual-world eye-tracking and gating methods to investigate how Spanish monolinguals and English late learners of Spanish use lexical stress to predict present-past verb tense suffixes in Spanish, and whether their proficiency and WM mediates their ability to do so. Lexical stress, henceforth stress, is the relative prominence of a syllable with respect to other syllables in a word. Although stress is contrastive in English and Spanish, Anglophones rely more on segmental information (vowel quality) to identify stress, whereas Hispanophones rely on suprasegmental information (pitch, duration, intensity).

This study is important to test L1 processing models, L2 processing models, and general cognition models. First, some L1 processing models maintain that suprasegmental integration is part of spoken word recognition, whereas others claim that integration and word recognition are separate mechanisms (see Roll, 2015, for a review), and little is known regarding the extent to which learners integrate cue weighting strategies of prosodic information in L2 lexical access. Second, some L2 processing models propose that late learners' difficulty processing inflectional morphology is due to their inability to integrate morphological cues to make predictions (Hopp, 2015), whereas others provide other representational and computational reasons (see Author, 2010, for a review). Third, some general cognition models argue that language processing is embedded in other aspects of our cognition (thus, memory influences language processing), whereas others defend that it is an isolated capacity (see Vogelzang et al., 2017, for a review).

Prosodic cues in L1 anticipation

According to top-down models of human information processing, people constantly and implicitly make probabilistic predictions based on previously encountered information and memory: from estimating how to move our hands when making a sandwich to predicting what a speaker is likely to say next (Bar, 2007; Clark, 2013). During sentence comprehension, anticipation is crucial because it allows for rapid and incremental integration of incoming information (Altmann & Mirkovic, 2009), and promotes efficient communication (Jaeger & Snider, 2013). Native speakers use various cues to predict non-linguistic information and linguistic information at all levels of representation (see Huettig et al., 2011; Kamide, 2008, for overviews). For example, there is robust evidence that monolinguals use morphological cues for morphosyntactic anticipation (e.g., Dahan, Swingley, Tanenhaus, & Magnuson, 2000; DeLong, Urbach, & Kutas, 2005; Lew-Williams & Fernald, 2007), and prosodic (specifically suprasegmental) cues for syntactic predictions (e.g., intonation for garden-path sentences, Steinhauer, Alter, & Friederici, 1999, and word order, Weber, Grice, & Matthew, 2006; tone for word order: Roll, Horne, & Lindgren, 2011; vowel duration for active-passive constructions, Rehrig, 2017; pauses for early- late closure sentences, Kjelgaard & Speer, 1999; pauses for syntactic constituency, Hawthorne & Gerken, 2014).

Prosody (i.e., intonation, tone, stress, rhythm) represents a crucial component to successfully understanding language because lexical activation is finely-tuned to the acoustic-phonetic detail of the speech signal. There is ample evidence that listeners use prosodic cues to predict syntactic structures (Cutler, et al., 1997) and recognize spoken words (McQueen, 2005). However, little is known about listeners' use of prosodic cues to anticipate grammatical

information within a word. Roll and colleagues report that Swedish monolinguals use word tone to pre-activate suffixes to predict nouns (Roll et al., 2010; Roll, 2015) and verbs (present and past tense suffixes: Roll, 2015; Söderström et al., 2012), as well as pseudowords and words with suffixes replaced with a cough (Söderström et al., 2017). The latter indicates that the prosody-morphology interaction is independent of the lexical status of the word stem, and that listeners can use prosody to fully recover suffixes they cannot hear. We investigate whether the findings of these tone studies carry over to other prosodic cues (word stress) and languages (Spanish).

Prosodic cues in L2 anticipation

Anticipation is thought to be an efficient and inevitable tool generally used by children (Borovsky, Elman, & Fernald, 2012) and adults (Huttig et al., 2011) during native-language processing. However, recent studies suggest that anticipation does not occur in all contexts (Nieuwland et al., 2017), or across all populations, such as children with specific language impairment (Andreu, Sanz-Torrent, & Trueswell, 2012), older adults with lower WM abilities (Federmeier, McLennan, Ochoa, & Kutas, 2002), and late L2 learners (discourse cues: Grüter, Rohde, & Schafer, 2017; syntactic cues: Grüter & Rohde, 2013). Relevant to our study, intermediate learners show limited or no integration of morphological cues to predict L2 morphosyntax (Dussias et al., 2013; Foucart et al., 2016; Foucart et al., 2014; Grüter et al., 2010; Hopp, 2015; Lew-Williams & Fernald, 2007), unless the two languages are typologically similar (beginners: Liburd, 2014; intermediates: Dussias et al. 2013; Foucart, et al., 2014). As for advanced learners, some studies show that they can use morphological cues for L2 morphosyntactic anticipation (Marull, 2017; for parallel results with L2 syntactic anticipation,

see Dussias et al., 2013; Kaan, Dallas, & Wijnen, 2010; Kaan, 2014), whereas others suggest they can't (Hopp, 2015; Martin et al., 2013; for similar results with L2 syntactic anticipation, see Rehrig, 2017).

Rehrig's dissertation is particularly relevant to our study because she examined prosodic cues (vowel duration). High intermediate Chinese late learners of English and English monolinguals listened to sentences with progressive active (*the boy was pushing the girl*) or passive verb stems (*the boy was pushed by the girl*), and chose one of two pictures (boy pushing girl, girl pushing boy). The eye-tracking data showed that the learners did not use the prosodic cues (longer vowel duration in monosyllabic than polysyllabic verbs); instead, they waited until the second NP to decide the syntactic structure. However, the lack of effects could be due to the learners not been proficient enough (proficiency was measured via subjective self-ratings instead of an objective test), or to the comparison of two syntactic constructions known to be learned early (active) and late (passive). We explored whether beginning and truly advanced English late learners of Spanish use other suprasegmental cues (stress) to form predictions within a word.

Individual differences in proficiency and WM and L2 anticipation

In the previous sections, we have discussed why anticipation is important during real-time processing, what cues are used to predict upcoming information, and what information is predicted. Another central question about anticipation in language processing is how prediction takes place, that is, which factors *drive* prediction, and which ones *mediate* prediction. Recently, scholars have demonstrated an increased interest in identifying prediction's driving factors, and have identified the following: probability distribution (cue reliability), a language production system used covertly, low-level associations, combinatorial mechanisms, event

knowledge, or a combination of the above (see Huettig, 2015, for an overview). By contrast, mediating factors such as processing speed, language proficiency, and WM have received little attention.

There is a vast literature indicating that L2 proficiency mediates L2 morphosyntactic processing (as proficiency increases so does sensitivity to morphosyntactic agreement violations) (see Author, 2010, for a review). However, L2 studies on proficiency and anticipatory language processing are virtually non-existent. Marull (2017) asked English-Spanish intermediate and advanced learners to complete a self-paced reading and a picture-selection task, and concluded that the advanced, but not the intermediate, learners used morphosyntactic cues to anticipate an upcoming word. WM, the activation and maintenance of short-lived memory items while performing sometimes complex and time-consuming cognitive tasks (Barrouillet & Camos, 2007), is essential for language processing (e.g., Christiansen & may also be important during anticipation. This is so because WM (a) determines what language is processed, (b) may interact with some and not other mechanisms of prediction (e.g., it may be more relevant for combinatorial than simple anticipatory mechanisms), and (c) can influence language-mediated anticipatory eye movements (i.e., WM may be important for language-vision interactions). Studies on WM and L2 morphosyntactic processing are inconclusive (see Author, 2016, for a review). Author explains these differences in terms of L2 proficiency and the cognitive load of both the task and the WM test. Recently, Huettig & Janse (2016) have examined WM and processing speed effects on Dutch monolinguals' use of morphological cues to predict morphosyntactic information, and they have found that enhanced WM and faster processing speed predict anticipatory eye movements. However, there are no L2 studies exploring a possible

association between WM and L2 anticipatory processing. We fill this gap by examining whether L2 proficiency and WM mediate late learners' integration of suprasegmental cues to pre-activate upcoming morphological information within a word. In particular, we focus on lexical stress.

Lexical stress in adult SLA

L2 research shows that native English speakers have trouble with both the production (Bullock & Lord, 2003; Lord, 2007) and perception (Face, 2000, 2005, 2006; Ortega Llebaria, Gu, & Fan, 2013; Saalfeld, 2012) of stress in Spanish. Various acoustic correlates are associated with stress, mainly F0 (Hz), duration (ms) and intensity (dB) (see Gordon & Roettger, 2017, for a comprehensive review). Stress can be variable or fixed. In English and Spanish, stress is variable and phonologically contrastive (e.g., Spanish: *Papa* 'potato' vs. *paPA* 'father;,' English: *PREsent* vs. *preSENT*), but these languages differ regarding stress in several ways. On one hand, English is considered a stress-timed language in which the intervals between stressed syllables have roughly the same duration. Prominence at the word level is mainly achieved through a process of vowel reduction. Unstressed vowels are often reduced to [ə], whereas stressed vowels typically maintain their spectral qualities, and, as a result, are perceived as being more prominent (Dauer, 1983; Hualde, 2005; Ladefoged, 2001). Spanish, on the other hand, is considered a syllable-timed language, i.e., each syllable has roughly the same duration and vowel quality, regardless of stress placement (Hualde, 2005; Colantoni, Steele, & Escudero, 2015). In this language, stressed syllables receive greater prominence via the aforementioned acoustic correlates (F0, duration, and intensity).

Given the similar phonological status of stress in Spanish and English, Anglophones' difficulties with this aspect of learning Spanish is initially puzzling. However, online studies on

the perception of stress by Spanish and English monolinguals demonstrate that Hispanophones and Anglophones process stress differently. Regarding Hispanophones, Soto-Faraco, Sebastián-Gallés, & Cutler (2001) asked Spanish monolinguals to listen to a prosodically matched (*prinCI-prinCIpio*) or mismatched (*PRINci-prinCIpio*) prime, and choose between two words presented on the screen (*PRINcipe-prinCIpio* ‘prince-beginning’). Compared to a control condition, the participants were faster after the matched prime and slower after the mismatched prime, indicating that, for Spanish listeners, stress is used to reduce competition in lexical access and plays a role equivalent to that of a speech segment in the processing and recognition of words. This does not appear to be the case for Anglophones. Cooper, Cutler, & Wales (2002) asked English monolinguals to complete the same cross-modal priming experiments used by Soto-Faraco et al. (2001) in English, with a matching (*ADmi-ADmiral*) or a mismatching (*admi-ADmiral*) prime preceding a word pair (*ADmiral-adMIration*). The results indicated that the matching condition fostered faster response times when compared to the control condition. Importantly, the mismatched condition did not slow down processing relative to the control, suggesting that English listeners have the ability to utilize stress when recognizing words, but do not do so in the same way as Spanish listeners. Specifically, for listeners of both languages, the investigators found an effect of match facilitation, but only the Spanish listeners showed mismatch inhibition.

Importantly, in Cooper et al. (2002) the first syllable of the target prime always received primary or secondary stress, but was never unstressed, precluding a direct comparison to Soto-Faraco et al. (2001). This limitation derives from the lack of minimal pairs in English based on stress. Those that do exist occur almost entirely in semantically related words (see Jensen,

1993). Spanish, for its part, contains numerous minimal stress pairs, both related and unrelated semantically. Thus, it behooves the Spanish listener to pay attention to stress, as it facilitates eliminating competing candidates in word recognition. This is not true for English listeners, as suprasegmental cues are likely redundant in the presence of segmental cues, i.e., stress has a weak functional load (Cutler, 2012; Tremblay, Bruersma & Couhlin, 2017). Therefore, while English listeners can perceive stress, evidence suggests it is unnecessary for lexical access because unstressed vowel reduction is sufficient for distinguishing word pairs. Furthermore, English word recognition seems to be unhindered when stress is purposely misplaced (Cutler, 1986).

In sum, the apparent difficulties of Anglophones with Spanish stress can likely be attributed to misplaced attention regarding the most relevant acoustic cues during initial stages of L2 learning. This shows that the acoustic correlates of stress—F0, duration, intensity, formant frequencies, etc.—can be weighted differently, i.e., the relative importance of these cues is language-specific (Chrabaszcz, Winn, Lin & Idsardi, 2014; Holt & Lotto, 2006). Native listeners focus on the relevant L1 acoustic cues in order to maximize efficiency, and naïvely use these same cues when learning an L2 (Ingvalson, Holt, & McClelland, 2011; Iverson et al., 2003). For this reason, L2 learners often have difficulties producing and perceiving prosodic information in the target language, as they have fine-tuned their processing abilities in a way that efficiently uses the acoustic cues that are most relevant in their L1. This begs the question as to whether native English speakers can learn to use prosodic information specific to Spanish in lexical access.

The current study

The present study aims to shed light on the role of suprasegmental cues on lexical access (whether suprasegmental cues are used to access words), age of acquisition (whether late learners can develop native-like anticipation processes using suprasegmental cues irrelevant in their L1 to identify stress), language proficiency (whether advanced learners exhibit more native-like patterns than beginning learners), and cognitive load (whether WM mediates the integration of suprasegmental cues to predict morphological information during lexical access). The specific research questions are:

1. *Do Spanish monolinguals use stress cues to pre-activate suffixes to make lexical predictions?* We predict that Spanish monolinguals will integrate stress cues to pre-activate morphological information that will allow them to make lexical predictions. This prediction follows L1 studies showing that Swedish monolinguals use word tone to anticipate the same type of morphological information we are investigating, namely present-past verb tense suffixes (Roll, 2015; Söderström et al., 2012).
2. *Do beginning and advanced English late learners of Spanish use stress cues to pre-activate suffixes to make lexical predictions?* We hypothesize that beginners will be ‘stress deaf,’ but that advanced learners will use stress cues to determine present-past verb differences before hearing the suffix. This hypothesis is based on online L2 studies showing proficiency effects on sensitivity to L2 morphosyntactic agreement violations (e.g., Author, 2010, 2011; Osterhout et al, 2008; Rossi et al., 2006), and on the use of morphosyntactic cues to predict morphological information (Marull, 2017). Considering that stress has a distinct functional load in English (Cooper et al, 2002) and Spanish

(Soto-Faraco et al., 2001), we expect the advanced learners to perform above chance but to benefit from stress cues to a lesser degree than the monolinguals.

3. *If they do, does WM influence Spanish monolinguals and English-Spanish late learners' ability to use stress cues to pre-activate suffixes to make lexical predictions? We anticipate that enhanced WM abilities will facilitate all participants' anticipatory eye movements and accuracy, provided that they use stress cues to predict morphological information. We base this predictions on studies showing WM effects on L1 anticipatory eye movements (Huettig & Jansen, 2016), and L2 morphosyntactic processing (Author, 2008, 2010; Faretta-Stutenberg, 2014; Havik et al., 2009; Keating, 2010).*

We used eye-tracking and gating methods to explore these issue. In visual-world eye-tracking tasks, participants look at images as they listen to linguistic stimuli (Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995), and fixations to the images reflect lexical activation as the acoustic signal unfolds (Allopenna, Magnuson, & Tanenhaus, 1998; Tanenhaus, Magnuson, Dahan, & Chambers, 2000). We chose eye-tracking because it provides more information about the time-course of prediction than ERPs (eye movements are continuously recorded, and allow one to detect prediction-driven fixations that occur much earlier than the predictable suffix). Finally, we included an offline gating task, in which participants only heard the verb stem, and had to guess the suffix, to assess the participants' ability to predict morphology in a context in which we can be sure they did not receive any acoustic information from the target suffix.

Method

Participants

There were 63 participants: 38 English learners of Spanish (12 beginners, 26 advanced), and 25 Spanish monolinguals. Data collection took place in a large North American university (learners) and the monolingual Spanish region of Teruel (Spanish controls). Participants were between 18 and 43 years old (WM and processing speed start decreasing around the age of 40: Park et al., 2003), right-handed (Oldfield, 1971), educated (they had at least completed high school education). Also, they had normal or corrected-to-normal vision and normal audition, completed all tasks, and received monetary compensation for participating.

The learners were English native speakers living in the United States who were born to English monolingual parents, belonged to a household and a community where English was the only language spoken, attended elementary/middle/high school and college in English, had not lived in a non-English-speaking country for more than three months, had no knowledge of other languages apart from English and Spanish, and began learning Spanish after the age of 12. Importantly, the learners had an average DELE score of 31.18 points; however, the distribution was bimodal. Thus we divided the learners into two groups: (1) late beginning learners (LB), with scores of 16.10 ± 3.95 standard deviations (*SD*) and (2) late advanced learners (LA), with scores of 46.27 ± 4.09 *SD*. A *t*-test for independent samples revealed that both groups scored statistically differently on the proficiency test: $t = 19.98$, $p < 0.001$. With regard to the Spanish monolinguals, they were born and raised in a monolingual community in Spain, had not lived in a bilingual community or abroad more than three months, had learned some English in school

starting between the ages of 8 and 12, and had no or minimal knowledge of other languages apart from Spanish and English.

Materials and Procedure

Participants completed the tests individually in one session in this order: a consent form, a language background questionnaire (5 minutes), an adapted version of the *Diploma de Español como Lengua Extranjera* (DELE) test (15-20 minutes; learners only), an eye-tracking task (20 minutes), a gating task (10 minutes), a WM task (10 minutes), a phonological short-term memory task (10 minutes, reported elsewhere), and a production task (15 minutes, reported elsewhere).

Language background questionnaire. It was administered in the participants' L1 and contained questions about age of acquisition, knowledge of other languages, experience living abroad, and language experience and use. The latter includes number of years they studied Spanish, location and length of time living abroad, contact hours per week with Spanish, and context (in school, at home, with friends, at work).

Spanish proficiency test. It was adapted from the *Diploma de Español como Lengua Extranjera* (DELE) exam, and had 56 multiple-choice items (36 testing grammar and 20 reading comprehension).

Eye-tracking task. The eye-tracker was an EyeLink 1000 Plus desktop mount machine from SR Research (sampling rate: 1k Hz; spatial resolution: .32° horizontal, 25° vertical; averaged calibration error: .01°). Participants were seated in front of a computer screen, and were asked to listen to sentences in Spanish and press a button as soon as they recognized one of the two words in the screen. After the instructions, participants placed their chin on a chin rest,

performed a nine-point calibration grid task, completed the practice trials, were given a chance to ask questions, and completed the rest of the experiment.

Each trial began with a drift correction, followed by a 250-ms blank screen. The visual stimuli (word pairs) appeared 1,000 ms before the onset of the oral stimuli (recorded sentence), following numerous visual-word eye-tracking studies. Participants listened to the sentence (e.g., *el director firma la factura* ‘the director signs the bill’), and pressed the right or left shift key as soon as they recognized which of the two words they had heard (see Figure 1).

INSERT FIGURE 1 HERE

They were instructed to press the key as soon as possible, and to not wait to hear the entire sentence. The program allowed participants to answer at or after the onset of the target word. Key presses before the onset of the target word did not stop the sound file and were not analyzed. After the participants pressed a key, there was a 500 ms blank screen, and the next trial began. No feedback was given during the experiment. Finally, images of words rather than objects were used based on a pilot eye-tracking study with Spanish monolinguals using the same oral stimuli with pictures. The results showed that they needed too much time to decode the images (sometimes they could not make a decision even after hearing the entire verb). This is in line with previous research showing that, in non-predictive contexts, phonological competitor effects are stronger with words than pictures (Huettig & McQueen, 2007; Ito, Dunn, & Pickerin, 2017).

Participants were randomly assigned to one of two versions. Each version contained one of the two conditions of a given word pair (e.g., if *firma* ‘(s)he signs’ appeared in version 1, then

firmó ‘he signed’ appeared in version 2). Both versions included a familiarization phase (practice trials, presented in a fixed pseudo-randomized order), and a testing phase (filler and experimental trials, presented using a Latin square design). A Latin square design distributed the filler and experimental sentences into blocks (each block containing fillers and only one sentence of a given condition), and the blocks appeared in a randomized order. The sentences within each block were pseudo-randomized to minimize the possibility that two experimental sentences of the same condition appeared next to each other.

There were 66 recorded sentences: 18 practice, 32 fillers, and 16 experimental (8 paroxytone, 8 oxytone). The practice sentences served to familiarize subjects with the task and the speaker’s acoustic characteristics, such as speech rate and vowel duration.

Visual stimuli. Each recorded sentence was paired with one of two word images (target, distractor). The two words appeared at each of the two sides (left, right) of the screen equally frequently in the practice and post-practice sentences. These word pairs had identical first syllables and focused on three contrasts: number (*col-coles* ‘cauliflower-cauliflowers’), lexical (*mar-marco* ‘sea-frame’), and (the experimental) tense (*firma-firmó* ‘he signs-signed’). Both the practice and the post-practice sentences had one third of each type of contrast. Importantly, word frequencies were determined with the dictionary of frequencies *LEXESP* (Sebastián-Gallés et al. 2000), and *t*-tests revealed no significant differences in frequency between the two conditions for any of the contrasts: number: $t = .364$, $p = .727$; lexical: $t = 1.222$, $p = .268$; tense (experimental): $t = 1.865$, $p = .082$. The experimental verbs were 2-syllables long, regular, transitive, and of the three possible conjugations in Spanish (*-ar*, *-er*, *-ir*). About half of the experimental verbs in the practice and in the post-practice trials contained a rhotic or nasal coda in the first syllable (CVC)

(*firma-firmó* ‘(s)he signs-signed’; *pinta-pintó* ‘(s)he paints-painted’), and half did not (CV) (*lava-lavó* ‘(s)he washes-washed’). We controlled for first syllable structure, because the longer syllable caused by the coda may be an additional cue to word prominence (see Face, 2001; 2005). Finally, the words were displayed in a BenQ XL2420TE display monitor at a resolution of 1920 x 1080 pixels.

Auditory stimuli. The recorded sentences contained words covered in basic Spanish classes, were 5-7 words long, and followed an SVO syntactic structure. In Spanish the SVO structure is used for non-neutral declarative statements presenting new information and is associated with a descending staircase intonational contour. The subjects and the objects of the sentences had 2-4 syllable count nouns, and the verbs were 2-3 syllables long. The subject nouns appeared once or twice, but the verbs and object nouns only once. Importantly, the verbs of the experimental sentences appeared in the middle of the sentence, to ensure consistency and avoid potential pre-linguistic processing or spillover effects.

The stimuli were recorded in a sound-attenuated booth. A Shure SM58 microphone captured the productions of a female native speaker of Peninsular Spanish. A Marantz Solid State Recorder PMD670 captured the utterances at a sampling rate of 44.1 kHz and 16-bit quantization. Each sentence was recorded three times (each time in a different pseudo-randomized order). The speaker was instructed to use a consistent speaking rate and standard intonation without overemphasizing any of the target words. The optimal iteration of the three repetitions was selected based on clarity. Praat (Boersma & Weenink, 2017) was used to normalize the volume to -18dB, and add 100 ms of leading and trailing silence. Afterwards, all recordings were inspected by hand in Praat to ensure they contained a descending staircase

intonational contour with progressively lower pitch accents anchored around tonic syllables (Roettger & Gordon, 2017), which is the common pattern found in standard Spanish (Hualde, 2005). The utterances were read with a speech rate of 3.03 ± 0.49 SD syllables per second. The average length of the sentences was 2.51 ± 0.22 SD seconds. Finally, participants used Sol Republic 1601-32 headphones.

Gating task. This task was adapted from Grosjean (1980), and was programmed and presented with E-Prime 2.0 Professional (Psychological Tools). Similar to the eye-tracking task, participants looked at a fixation sign followed by a 250 ms blank screen, read two words, 1,000 ms later listened to a Spanish sentence, and pressed a right or a left button as soon as they were able to guess the word. They received no feedback. Different from the eye-tracking task, the target words of the non-practice sentences of the gating task were segmented at the offset of the first syllable, and participants had to guess what the whole word might be without hearing the suffix.

Participants listened to 114 sentences: 18 practice sentences with whole target words (e.g., *el mecánico busca las llaves* ‘the mechanic looks for the keys’), and 96 non-practice sentences (64 fillers, 32 experimental) with cut target words (e.g., *la persona dice firm* ‘the person says (s)he sign’ for the words *firma-firmó* ‘(s)he signs-signed’). Participants were warned that, after the practice sentences, they would listen to sentences ending in cut words, and that they would need to make a guess as soon as possible. Because the gating task made the target words obvious and this explicitness could bias the results of the eye-tracking task, the gating task was administered after the eye-tracking task. The filler and experimental sentences had the same sentence beginning (*la persona dice* ‘the person says’), and the target words were identical to

those of the eye-tracking task for comparability purposes. Finally, the sound files were recorded by the same speaker under the same conditions as in the eye-tracking task.

Working memory task. It was adapted from the letter-number sequencing test within the Wechsler Adult Intelligence Scale test (WAIS) (Wechsler, 1997). It consisted of 2 practice trials and 21 experimental trials ranging from 2 to 9 letter-number combinations. For each combination, participants saw a 500-ms fixation sign (+), followed by random letters and numbers presented non-cumulatively one at a time (e.g., 7-J-M-3). When the word RECALL appeared on the screen, they typed first the numbers in ascending order and then the letters in alphabetical order (e.g., 37JM). A non-linguistic test of verbal WM was employed to avoid confounding results caused by Spanish proficiency shortcomings. One point was awarded per set if all the letters and numbers were recalled in the correct order.

Results

Eye-tracking task

The eye tracking data were downsampled to 10 and 50 ms bins, following previous psycholinguistic studies (e.g., Allopenna et al., 1998), and time windows were subsequently shifted forward 200 ms, which represents the minimum time necessary to plan and launch a saccade (e.g., Fischer, 1992; Matin, Shao, & Boff, 1993; Saslow, 1963). We then submitted the data to three analyses. The objective of the first analysis was to determine whether or not the participants were able to predict the morphology of the target word above chance. To this end, we centered the time course for all trials around the 10 ms bin that occurred 200 ms after the offset of the first syllable of the target word. Next, we calculated by-subject target fixations and submitted the scores to one-tailed t-tests for each group, for each level of stress (paroxytone,

oxytone). The purpose of this analysis was to determine whether or not the group target fixations were above chance after hearing the first syllable of the target word. We set μ at 50% and the alternative hypothesis was that target fixations were greater than μ . Alpha was bonferroni corrected ($0.05 / 6 = 0.008$) to avoid family-wise error.

The analysis revealed that only the native speakers fixated on the target words above chance at the offset of the target syllable. This was true for both the paroxytone ($t(21) = 2.80$, Estimate = 0.63, CI low = 0.55, $p < 0.005$) and oxytone ($t(21) = 4.77$, Estimate = 0.72, CI low = 0.64, $p < 0.001$) items. Neither of the learner groups fixated on the targets above chance at our a priori adjusted alpha level (See Table 3 for a complete summary).

INSERT TABLE 1 HERE

The second analysis of the eye-tracking data examined target fixations as a function of the fixed effects group (SS, LA, LB), stress (oxytone, paroxytone), syllable structure (CV, CVC) and standardized working memory scores. SS stands for native Spanish speakers, LA for late advanced learners, and LB for late beginning learners. Again, this analysis utilized the data from the 10 ms bin centered around the offset of the first syllable of target items. Due to the categorical nature of the outcome variable (target, competitor), the data were analyzed using GLMMs with a binomial linking function. The categorical fixed effect group was dummy coded with SS set as the baseline. The fixed effects stress and syllable structure were deviation coded such that the omnibus model provided an assessment of main effects. The random effects

structure included by-subject and by-item random intercepts with random slopes for stress and syllable structure.

For all multilevel regression models main effects and interactions were assessed by hierarchically partitioning the variance via nested model comparisons. Alpha was set at 0.05. All statistical analyses were carried out using R (R Core Team, 2017). We used *lme4* (Douglas et al., 2015) to fit the mixed effects models and *lsmeans* (Lenth, 2016) for multiple comparisons. *MuMIn* (Bartoń, 2016) provided an assessment of the variance explained by the models with the random effects structure (conditional R^2 or R^2_c) and without (marginal R^2 or R^2_m) (See Nakagawa & Schielzeth, 2013).

The GLMM was best fit when including the random effects ($R^2_m = 0.05$, $R^2_c = 0.31$). There was a main effect of group ($\chi^2(2) = 17.12$, $p < 0.001$) and syllable structure ($\chi^2(1) = 7.57$, $p < 0.006$), but not of stress ($\chi^2(1) = 0.05$, $p > 0.05$) or WM ($\chi^2(1) = 1.15$, $p > 0.05$), nor were there any higher order interactions. Both of the learner groups were less likely to fixate on the target than the native controls. For the LA group the log odds of fixating on the target were reduced by 1.25 ± 0.34 se ($z = -3.62$, $p < 0.001$), and for the LB group the log odds were reduced by 1.52 ± 0.39 se ($z = -3.93$, $p < 0.001$). When comparing the learner groups directly, there was no change in the log odds of fixating on the target ($p > 0.05$). Overall, participants were less likely to fixate on the target items that did not have a heavy initial syllable ($\beta = -0.53$, $SE = 0.19$, $z = -2.79$, $p < 0.006$). Figure 3 illustrates the final model output.

INSERT FIGURE 2 HERE

Finally, we used Growth Curve Analysis (Mirman, 2014) to analyze the time course of fixation from 500 ms before the offset of the initial syllable of the target word to 250 ms after (i.e., from the average target word onset to when the target fixations plateaued). We applied the empirical logit transformation (Barr, 2008) to the outcome variable (fixating on the target or not). To apply the transformation, we down-sampled the time course data into 50 ms bins. Next, we summed target and distractor fixations for each bin and calculated the log-odds of looking to the target word after adding 0.5 to both the numerator and denominator as to avoid dividing by 0. Each bin is then weighted using the formula given in Equation 1:

$$w(Y, N) = \frac{1}{Y+0.5} + \frac{1}{N-Y+0.5} \quad (1)$$

Where Y equals the summed target looks and N equals the number of trials. The empirical logit transformation is often used in order to avoid boundary condition problems. It is also helpful in increasing computational speed, as well as allowing for more complex random effects structures (See Barr, 2008; and Mirman, 2014 for discussion). We modeled the time course using linear, quadratic and cubic orthogonal polynomials, with fixed effects of group (SS, LA, LB), stress (paroxytone, oxytone), syllable structure (CV, CVC) and standardized working memory on all time terms. Group was coded using the native speakers as the baseline, thus the LB and LA parameters described how the growth curve of the learners differed from the native controls. Stress and syllabic structure were deviation coded. All models included participant and participant-by-condition random effects on all time terms. Main effects and higher order interactions were assessed using model comparisons.

The overall time course data are shown in Figure 4. One can observe that target fixations increased as a function of time. Specifically, the three groups fixated on the target at chance from the onset of the sentence up to around 250 ms into the target word. The proportion of fixations then increased around the offset of the first syllable of the target word and plateaued approximately 250 ms afterwards.

INSERT FIGURE 3 HERE

Model estimates are plotted in Figure 5. There were no significant effects associated with the cubic terms, stress, or WM, therefore they were disregarded from further analyses. There was a main effect of group on the intercept ($\chi^2(2) = 10.52, p < 0.006$) and the quadratic term ($\chi^2(2) = 10.72, p < 0.005$), thus the grouping variable and the higher order polynomials were retained for model comparisons. With regard to the LA group, the intercept suggests that the SS group had a slightly higher overall fixation probability (i.e., more area under the curve), though this effect was not lower than our specified alpha (Estimate = -0.37, SE = 0.20, $p = 0.06$). The significant quadratic term shows that the SS group was quicker to recognize the target than the LA group (Estimate = 1.07, SE = 0.32, $p < 0.002$). With regard to the LB group, the intercept (Estimate = -0.57, SE = 0.22, $p < 0.02$), and the quadratic term (Estimate = 1.04, SE = 0.35, $p < 0.004$) were significant, indicating that the SS group had a higher overall fixation probability, and faster recognition of the target. The positive value of both quadratic terms indicates that the growth curves for the learner groups were bowed in the opposite direction when compared to the native controls, revealing slower target recognition over the time course. There was also an effect of

syllable structure on the intercept of the LA group (Estimate = 0.18, SE = 0.07, $p < 0.02$), indicating a higher overall fixation probability in the presence of a coda. This is observable in Figures 4 and 5, as the green line appears to be phase-shifted to the left in the CVC condition. There were no other significant effects. The complete output of the final model is available in Appendix A.

INSERT FIGURE 4 HERE

Gating task

The gating data were submitted to two different statistical analyses. The first analysis paralleled that of the t-tests conducted with the eye-tracking data. Thus, we calculated by-subject response accuracy and submitted the scores to one-tailed t-tests for each group, in each stress condition (paroxytone, oxytone). We set μ at 50% accuracy, and the alternative hypothesis was that scores were greater than μ . Moreover, we bonferroni corrected alpha ($0.05 / 6 = 0.008$) to avoid capitalizing on chance due to family-wise error.

The one-tailed t-tests indicated that the LA and SS groups were able to accurately predict the morphology of the target items in the paroxytone condition (SS: $t(24) = 7.56$, Estimate = 0.76, CI low = 0.70, $p < 0.001$; LA: $t(25) = 6.71$, Estimate = 0.77, CI low = 0.70, $p < 0.001$), and the oxytone condition (SS: $t(24) = 7.02$, Estimate = 0.76, CI low = 0.69, $p < 0.001$; LA: $t(25) = 8.87$, Estimate = 0.79, CI low = 0.73, $p < 0.001$). The LB group, on the other hand, did not predict the target morphology above chance for either condition (Paroxytone: $t(9) = 2.15$,

Estimate = 0.65, CI low = 0.52, $p > 0.008$; Oxytone: $t(9) = 0.51$, Estimate = 0.51, CI low = 0.38, $p > 0.008$). It is worth noting that the LB group was more accurate responding to paroxytone targets (Mean: 0.65, SD = 0.48) than oxytone targets (Mean = 0.51, SD = 0.50), however, not significantly above chance at our adjusted alpha level. Table 2 summarizes the results of the t-tests.

INSERT TABLE 2 HERE

The objective of the second analysis of the gating data was to determine if response accuracy varied as a function of the fixed effects group (SS, LA, LB), stress (oxytone, paroxytone), syllable structure (CV, CVC) and standardized WM scores. Given the categorical nature of the outcome variable (correct, incorrect), the data were analyzed using generalized linear mixed effects models (GLMMs) with a binomial linking function. We utilized the same coding schemes and random effects structure described in the eye-tracking experiment. Likewise, main effects and higher order interactions were assessed by hierarchically partitioning the variance via nested model comparisons with alpha set at 0.05.

The GLMM was best fit with the maximal error structure ($R^2_m = 0.03$, $R^2_c = 0.27$). The model yielded a main effect of group ($\chi^2(2) = 19.32$, $p < 0.001$) and there was a marginal effect of working memory ($\chi^2(1) = 2.75$, $p = 0.09$). There were no main effects of stress ($\chi^2(1) = 0.08$, $p > 0.05$) or syllable structure ($\chi^2(1) = 0.07$, $p > 0.05$), nor were there any higher order interactions. Specifically, when comparing the SS group to the LB group, the log odds of responding correctly were reduced by 1.06 ± 0.24 se ($z = -4.49$, $p < 0.001$) independent of stress and the presence or

absence of a coda. The likelihood of responding correctly did not differ between the SS and LA groups ($p > 0.05$). The LB group was also less likely to respond correctly than the LA group ($\beta = 1.11$, $SE = 0.21$, $z = 4.61$, $p < 0.001$). In sum, the SS and LA groups did not differ from each other in their response behavior for the gating task, but the LB group was less accurate than both the advanced learners and the native speakers. For all groups, accuracy was not affected by stress, or syllabic structure. Regarding the marginal effect of WM, the model suggests that increased working memory might be associated with higher accuracy ($\beta = 0.14$, $SE = 0.08$, $z = 1.69$, $p = 0.09$), however, this effect should be interpreted with caution.

INSERT TABLE 3 HERE

INSERT FIGURE 5 HERE

Discussion

We next discuss the theoretical implications of our findings.

1. Do Spanish monolinguals use stress cues to anticipate suffixes to make lexical predictions?

Our prediction that the Spanish monolinguals would use stress information to predict words was supported. Thus, the three analyses of the eye-tracking data show that monolinguals fixated on the target words before they were segmentally distinguishable, and the results of the

gating task reveal that they accurately predicted the target suffixes. Importantly, monolinguals successfully predicted the target words above chance in both tasks, regardless of stress (paroxytone, oxytone) and syllabic structure (CVC, CV).

These findings are in line with Roll and colleagues' experiments showing that Swedish monolinguals use word tone to predict present-past tense verb suffixes (Roll, 2015; Söderström et al., 2012). Our study contributes to this body of literature by demonstrating that these findings extend to other languages (Spanish) and suprasegmental cues (stress), as well as to L2 learning. Importantly, other experiments with L1 Swedish and word tone report similar findings when predicting nouns (Roll et al., 2010; Roll, 2015), as well as pseudowords and words with suffixes replaced with a cough (Söderström et al., 2017). Taken together, the findings of these studies and our study indicate that listeners predict probabilistically, that they use suprasegmental cues to anticipate morphological information and make lexical predictions, and that the the prosody-morphology interaction is independent of the lexical status of the word stem. This poses a challenge to models maintaining that structural integration and word recognition are separate mechanisms.

2. Do beginning and advanced English learners of Spanish use stress cues to anticipate suffixes to make lexical predictions?

Our hypotheses that advanced learners, but not beginners, would use stress cues to determine present-past verb differences before hearing the suffix, and that advanced learners would use stress less than the monolinguals were supported. While beginners were sensitive to stress, they did not fixate on targets until after verb suffixes, suggesting that prosodic information is not used in early stages of learning for anticipatory purposes. This is likely explained by the

functional load of stress in English versus Spanish. As for advanced learners, the gating task showed that advanced learners were as accurate as monolinguals regardless of stress and syllabic structure. However, the same was not true in the eye-tracking task, where the advanced learners only anticipated word morphology if the target word contained extra acoustic information. In this case a nasal coda.

Previous research has shown that syllable weight affects the perception of stress in L1 and L2 listeners. Specifically, heavy syllables, those with a long vowel or a coda, are perceived as being more prominent when pitch, duration and intensity are controlled (see Face, 2001, 2005). Thus, the combination of the heavy syllable and the acoustic cues of stress in these words may have provided the listener with sufficient acoustic information, which appears to be necessary in an online task such as eye-tracking. Why, then, were the advanced learners accurate with and without the coda in the gating task? This difference is likely explained by the nature of the task. We analyzed response accuracy, an offline measurement, thus the participants had time to consider their responses.

The results of our second research question are in line with L2 anticipation studies showing integration of morphological cues to predict L2 morphosyntax at advanced high proficiency levels (Marull, 2017; for similar results with L2 syntactic anticipation, see Dussias et al, 2013; Kaan et al, 2013), but reduced or no integration at low proficiency levels (Dussias et al., 2013; Foucart et al., 2016; Foucart et al., 2014; Grüter et al., 2010; Hopp, 2015; Lew-Williams & Fernald, 2010; Marull, 2017), unless the two languages are typologically similar (beginners: Liburd, 2014; intermediates: Dussias et al. 2013; Foucart, et al., 2014).

In addition to informing lexical access models by showing that adults can integrate L2 suprasegmental cues used differently in the L1 (accessibility accounts) for lexical anticipation, the findings of our second research question shed light on current top-down models of L2 processing that relate late learners' persistent difficulty acquiring inflectional morphology to problems integrating grammatical information during real-time processing. According to these models, L2 learners' integration problems are caused by L1-L2 neurocognitive differences (e.g., Weber-Fox & Neville, 1996; Ullman, 2005), limited WM resources (e.g., Author, 2010; McDonald, 2006), inability to encode morphosyntactic information as implicit knowledge (e.g., Clahsen & Felser, 2006; Jiang, 2004, 2007), or failure to use morphosyntactic cues to anticipate upcoming words (e.g., Hopp, 2015). Importantly, studies on L2 morphosyntactic anticipation, processing, and production concentrate on suffixes, supposedly the root of the problem. We propose additional explanations to late learners' difficulty processing inflectional morphology in terms of their impoverished abilities to integrate prosodic information to make predictions. Based on our findings, we argue that adult learners can gain anticipatory patterns qualitatively similar to monolinguals (both use stress to anticipate suffixes), but they behave quantitatively different from monolinguals (monolinguals can anticipate the suffixes earlier). A similar study with near-native learners would determine if late learners can make predictions as fast as monolinguals.

3. Does WM influence Spanish monolinguals and English-Spanish late learners' ability to use stress to anticipate suffixes to make lexical predictions?

Our expectations that enhanced WM abilities would facilitate the participants' anticipatory eye movements and accuracy based on Huettig and Jansen's (2016) L1 findings was

partially supported. In the eye-tracking task, WM was not associated with earlier target fixations or accuracy for any of the three groups. But in the gating task, there was a marginal main effect of WM on accuracy, such that higher WM abilities yielded higher accuracy scores. However, we report this finding with caution because the probability of our data, given our hypothesis, was not significant at our more stringent corrected alpha level.

The WM effects in the gating task, but not the eye-tracking task, are likely explained by the eye-tracking task being online, whereas the gating task was offline. Therefore, in an offline task, when given time and ample acoustic information, the advanced learners perform similarly to natives. In an online task, we can see that the extra acoustic information speeds processing. In its absence, target fixations occur later in the time course. In sum, it appears that added acoustic information by way of syllabic structure can increase accuracy, but at this juncture, the role of WM regarding the ability to predict morphology is unclear.

Conclusion

The present study tested whether Spanish monolinguals and late L2 learners use suprasegmental cues to pre-active morphological information during spoken word recognition, and if they do, whether this ability is modulated by proficiency and WM. We investigated these issues with visual-world eye-tracking and gating methods. In both tasks, participants saw two words on a screen with (*firma* ‘(s)he signs’) and without (*firmÓ* ‘(s)he signed’) word initial stress, heard a sentence containing one of the two words, and chose the word they had heard. Eye-tracking sentences contained complete target words (*el director firma la factura* ‘the director signs the bill’) (eye-tracking task), but gating sentences had target words missing the suffix (*la persona dice: firm* ‘the person says: (s)he sign’). The results showed that the

monolinguals looked at the target word before it was segmentally distinguishable in the time course regardless of the task (eye-tracking, gating), word initial stress (paroxytone, oxytone), and syllabic structure (CV, CVC first syllable). The same was true for the advanced learners for all the verbs of the gating task, and for CVC, but not CV, verbs of the eye-tracking task, although they were slower than the monolinguals making predictions. In contrast, the beginners failed to use stress to make lexical predictions in the eye-tracking task and the gating task. Finally, WM was not associated with earlier target fixations for any of the three groups in the eye-tracking task, but higher WM was related to higher accuracy in the gating task (marginal main effect) probably to the offline nature of this task. Taken together, these results suggest that suprasegmental information about stress can guide lexical access in monolinguals and adult learners, and that this skill develops with proficiency. We argue that in L2 acquisition the weighting of segmental and suprasegmental cues of the target language can be learned and effectively computed during lexical access. We also propose that late learners' difficulty acquiring L2 inflectional morphology may be due to their impoverished abilities to integrate prosodic cues that are less reliable in the L1 than the L2 to pre-activate morphological information. Finally, the findings of this study advance our understanding of the role of anticipation in the human brain (Huetting, 2015).

References

- Allopenna P. D., Magnuson J. S., & Tanenhaus M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language*, 38, 419–439.
- Altmann, G. T. & Mirkovic, J. (2009). Incrementality and prediction in human sentence processing. *Cognitive Sciences*, 33, 583-609.
- Andreu, L., Sanz-Torrent, M., & Trueswell, J. C. (2012). Anticipatory sentence processing in children with specific language impairment: Evidence from eye movements during listening. *Applied Psycholinguistics*, 34, 5-44.
- Author (2010)
- Author (2016)
- Barr, D. J. (2008). Analyzing “visual world” eyetracking data using multilevel logistic regression. *Journal of Memory and Language*, 59(4), 457–474.
- Bar, M. (2007). The proactive brain: Using analogies and associations to generate predictions. *Trends in Cognitive Sciences*, 11(7), 280–289.
- Bartoń, K. (2016). MuMIn: Multi-Model Inference. R package version 1.15.6.
- Barrouillet, P. & Camos, V. (2007). The time-based resource-sharing model of working memory. In N. Osaka, R. H. Logie, & M. D’Esposito (Eds.), *The cognitive neuroscience of working memory* (pp. 59-80). Oxford: Oxford University Press.
- Boersma, Paul & Weenink, David (2017). Praat: doing phonetics by computer [Computer program]. Version 6.0.33, retrieved 26 September 2017 from <http://www.praat.org/>

- Borovsky, A., Elman, J. L., & Fernald, A. (2012). Knowing a lot for one's age: Vocabulary skill and not age is associated with anticipatory sentence interpretation in children and adults. *Journal of Experimental Child Psychology*, 112, 417-436.
- Bullock, B., & Lord, G. (2003). Analogy as a learning tool in second language acquisition: The case of Spanish stress. *Amsterdam studies in the theory and history of linguistic science*, 281-297.
- Chrabaszcz, A., Winn, M., Lin, C. Y., & Idsardi, W. J. (2014). Acoustic cues to perception of word stress by English, Mandarin, and Russian speakers. *Journal of Speech, Language, and Hearing Research*, 57(4), 1468-1479.
- Christiansen, M., & Chater, N. (2016). The now-or-never bottleneck: a fundamental constraint on language. *Brain Behavioral Sciences*, 39, e2.
- Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral Brain Sciences*, 36, 181-204.
- Clahsen, H. & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics*, 27, 3-42.
- Colantoni, L., Steele, J., Escudero, P. (2015). *Second Language Speech: Theory and Practice*. Cambridge: Cambridge University Press.
- Cooper, N., Cutler, A., & Wales, R. (2002). Constraints of lexical stress on lexical access in English: Evidence from native and non-native listeners. *Language and speech*, 45(3), 207-228.
- Cutler, A. (1986). Forbear is a homophone: Lexical prosody does not constrain lexical access. *Language and Speech*, 29, 201-220.

- Cutler, A. (2012). *Native listening: Language experience and the recognition of spoken words*. Cambridge, MA: MIT Press.
- Cutler, A., Dahan, D., & van Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech*, 40(2), 141-201.
- Dahan, D., Swingle, D., Tanenhaus, M. K., & Magnuson, J. S. (2000). Linguistic gender and spoken-word recognition in French. *Journal of Memory and Language*, 42, 465-480.
- Dauer, R.M. (1983). Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics*, 11, 51-62.
- Dell, G. S. & Chang, F. (2014). The P-chain: relating sentence production and its disorders to comprehension and acquisition. *Philosophical Transactions of the Royal Society London Series B, Biological Sciences*, 369, 20120394.
- DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience*, 8, 1117-1121.
- Douglas, B., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1-48.
- Dussias, E. P., Valdés Kroff, J. R., Guzzardo Tamargo, R. E., & Gerfen, C. (2013). When gender and looking go hand in hand. Grammatical gender processing in L2 Spanish. *Studies in Second Language Acquisition*, 35, 353-387.
- Face, T. L. (2000). The role of syllable weight in the perception of Spanish stress. In H. Campos, E. Herburger, A. Morales-Front, & T. J. Walsh (Eds.), *Hispanic linguistics at the turn of the millennium* (pp. 1-13). Somerville, MA: Cascadilla Proceedings Project.

- Face, T. L. (2005). Syllable weight and the perception of Spanish stress placement by second language learners. *Journal of Language and Learning*, 3(1), 90-103.
- Face, T. L. (2006). Cognitive factors in the perception of Spanish stress placement: implications for a model of speech perception. *Linguistics* 44–6 (2006), 1237–1267.
- Federmeier, K., McLennan, D. B., De Ochoa, E., & Kutas, M. (2002). The impact of semantic memory organization and sentence context information on spoken language processing by younger and older adults: An ERP study. *Psychophysiology*, 39, 133-146.
- Fischer, B. (1992). Saccadic reaction time: Implications for reading, dyslexia and visual cognition. In K. Rayner (Ed.), *Eye movements and visual cognition: Scene perception and reading* (pp. 31-45). New York: Springer Verlag.
- Foucart, A., Martin, C., Moreno, E. M., & Costa, A. (2014). Can bilinguals see it coming? Word anticipation in L2 sentence reading. *Journal of Experimental Psychology*, 40(5), 1461-9.
- Foucart, A., Ruiz-Tada, E., & Costa, A. (2016). Anticipation processes in L2 speech comprehension: Evidence from ERPs and lexical recognition tasks. *Bilingualism: Language and Cognition*, 19(1), 213-219.
- Gordon, M., & Roettger, T. (2017). Acoustic correlates of word stress: A cross-linguistic survey. *Linguistics Vanguard*, 3(1).
- Grosjean, F. (1980). Spoken word recognition processes and the gating paradigm. *Perception and Psychophysics*, 28(4), 267-283.
- Grüter, R., Lew-Williams, C., & Fernald, A. (2012). Grammatical gender in L2: A production or a real-time processing problem? *Second Language Research*, 28(2), 191-215.

- Grüter, T., & Rohde, H. (2013). L2 processing is affected by RAGE: Evidence from reference resolution. Talk presented at the 12th Conference on Generative Approaches to Second Language Acquisition (GASLA). University of Florida, FL.
- Grüter, T., & Rohde, H., & Schafer (2017). Coreference and discourse coherence in L2. The roles of grammatical aspect and referential form. *Linguistic Approaches to Bilingualism*, 7(2), 199-229.
- Hawthorne, K. & Gerken, L. A. (2014). From pauses to clauses: Prosody facilitates learning of syntactic constituency. *Cognition*, 133(2), 420-428.
- Holt, L. L., & Lotto, A. J. (2006). Cue weighting in auditory categorization: Implications for first and second language acquisition. *The Journal of the Acoustical Society of America*, 119, 3059-3071.
- Hopp, H. (2015). Semantics and morphosyntax in predictive L2 sentence processing. *IRAL*, 53(3), 277-306.
- Hualde, J. I. (2005). *The sounds of Spanish*. New York City: Cambridge University Press.
- Huetting, F. (2015). Four central questions about prediction in language processing. *Brain Research*, 1626, 118-135.
- Huetting, F., and Janse, E. (2016). Individual differences in working memory and processing speed predict anticipatory spoken language processing in the visual world. *Lang. Cogn. Neurosci.* 31, 80–93.
- Huetting, F., & McQueen, J. M. (2007). The tug of war between phonological, semantic and shape information in language-mediated visual search. *Journal of Memory and Language*, 57(4), 460-482.

- Huetting, F., Rommers, J., & Meyer, A. S. (2011). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta Psychologica, 137*(2), 151-171.
- Ingvalson, E. M., Holt, L. L., & McClelland, J. L. (2011). Can native Japanese listeners learn to differentiate /r-l/ on the basis of F3 onset frequency? *Bilingualism: Language and Cognition, 15*, 255-274.
- Iverson, P., Kuhl, P. K., Akahane-Yamada, R., Diesch, E., Tohkura, Y. i., Kettermann, A., & Siebert, C. (2003). A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition, 87*, B47-B57.
- Ito, A., Dunn, M. S., & Pickering, M. (2017). Effects of language production on prediction: Word vs. picture visual world study. Poster presented at *Architectures and mechanisms of language processing*.
- Jaeger, T. F & Snider, N. E. (2013). Alignment as a consequence of expectation adaptation: Syntactic priming is affected by the prime's prediction error given both prior and recent experience. *Cognition, 127*(1), 57-83.
- Jensen, J. T. (1993). *English phonology*. Philadelphia: John Benjamins.
- Jiang, N. (2004). Morphological insensitivity in second language processing. *Applied Psycholinguistics, 25*, 603-634.
- Jiang, N. (2007). Selective integration of linguistic knowledge in adult second language learning. *Language Learning, 57*(1), 1-33.

- Kaan, E., Dallas, A., Wijnen, N.K. (2010). Syntactic predictions in second-language sentence processing. In: Jan-Wouter Zwart and Mark de Vries (eds) . Structure preserved. Festschrift in the honor of Jan Koster. John Benjamins. (pp. 208-213).
- Kaan, E. (2014). Predictive sentence processing in L1 and L2. What's different? *Linguistic Approaches to Bilingualism*, 4(2), 257-282.
- Kamide, Y. (2008). Anticipatory processes in sentence processing. *Language and Linguistics Compass*, 2(4), 647-670.
- Kjelgaard, M., & Speer, S. (1999). Prosodic facilitation and interference in the resolution of temporary syntactic closure ambiguity. *Journal of Memory and Language*, 40, 153–194.
- Ladefoged, P. (2001). *Vowels and consonants: an introduction to the sounds of languages*. Oxford: Blackwell.
- Lenth, R. V. (2016). Least-Squares Means: The R Package lsmeans. *Journal of Statistical Software*, 69(1), 1-33.
- Lew-Williams, C. & Fernald, A. (2007). Real-time processing of gender-marked articles by native and non-native Spanish speakers. *Journal of Memory and Language*, 63(4), 447-464.
- Liburd, T. (2014). Investigating prediction in L2 morpho-syntax: A visual world study. Unpublished doctoral dissertation. University of Pittsburgh.
- Lord, G. (2007). The role of lexicon in learning second language stress patterns. *Applied Language Learning*, 17(1-2), 1-14.
- Marull, C. (2017). Second language processing efficiency: Experience and cognitive effects on L2 morphosyntactic integration and anticipation. In M. LaMendola & J. Scott (Eds.),

- Proceedings of the 41st Boston University Conference on Language Development (BUCLD)* (pp. 466-480).
- Martin, C. D. et al., (2013). Bilinguals reading in their second language do not predict upcoming words as natives do. *Journal of Memory and Language*, 69, 574-588.
- Matin, E., K.C. Shao, & Boff, K.R. (1993). Saccadic overhead: Information-processing time with and without saccades. *Perception and Psychophysics*, 53, 372–380.
- McDonald, J. (2006). Beyond the critical period: Processing-based explanations for poor grammaticality judgment performance by late second language learners. *Journal of Memory and Language*, 55, 381-401.
- McQueen, J. M. (2005). Spoken word recognition and production: Regular but not inseparable bedfellows. In A. Cutler (Ed.), *Twenty-first century psycholinguistics: Four cornerstones* (pp. 229-244). Mahwah, NJ: Erlbaum.
- Mirman, D. (2014). *Growth curve analysis and visualization using R*. Boca Raton, FL: Chapman & Hall/CRC.
- Nakagawa, S. & Schielzeth H. (2013) A general and simple method for obtaining R² from generalized linear mixed-effects models. *Methods in Ecology and Evolution* 4: 133–42.
- Nieuwland, M. S., et al. (2017). Limits on prediction in language comprehension: A multi-lab failure to replicate evidence for probabilistic pre-activation of phonology. *BioRxiv*. doi: <https://doi.org/10.1101/111807>
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, 9, 97-113.

- Ortega Llebaria, M., Gu, H., & Fan, J. (2013). English speakers' perception of Spanish lexical stress: Context-driven L2 stress perception. *Journal of Phonetics* 41 (2013) 186–197.
- Park, D., Welsh, R., Marschuetz, C., Gutchess, A., Mikels, J., Polk, T., Noll, D., & Taylor, S. (2003). Working memory for complex scenes: Age differences in frontal and hippocampal activations. *Journal of Cognitive Neuroscience*, 15(8), 1122-1134.
- Pickering, M. J. & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral Brain Sciences*, 36, 329-347.
- R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Rehrig, G. L. (2017). Acoustic correlates of syntax in sentence production and comprehension. Unpublished doctoral dissertation. Rutgers University.
- Roettger, T., & Gordon, M. (2017). Methodological issues in the study of word stress correlates. *Linguistics Vanguard*, 3(1).
- Roll, M. (2015). A neurolinguistic study of South Swedish word accents : Electrical brain potentials in nouns and verbs. *Nordic Journal of Linguistics*, 38(2), 149–162.
- Roll, M., Horne, M., & Lindgren, M. (2010). Word accents and morphology - ERPs of Swedish word processing. *Brain Research*, 1330, 114-123.
- Roll, M., Horne, M., & Lindgren, M. (2009). Left-edge boundary tone and main clause verb effects on syntactic processing in embedded clauses - an ERP study. *Journal of Neurolinguistics*, 22, 55–73.
- Roll, M., Söderström, P., & Horne, M. (2013). Word-stem tones cue suffixes in the brain. *Brain Research*, 1520, 116-120.

- Saalfeld, A. K. (2012). Teaching L2 spanish stress. *Foreign language annals*, 45(2), 283-303.
- Saslow, M. G. (1967). Latency for saccadic eye movement. *Journal of the Optical Society of America*, 57, 1030-1033.
- Sebastián-Gallés, N., Martí, M. A., Cuetos, F., & Carreiras, M. F. (2000). *LEXESP: Léxico informatizado del español [LEXESP: A computerized word-pool in Spanish]*. Barcelona: Edicions de la Universitat de Barcelona.
- Söderström, Horne, M., & P., Roll, M. (2017). Stem tones pre-activate suffixes in the brain. *Journal of Psycholinguistic Research*, 46(2), 271-280.
- Söderström, P., Roll, M., & Horne, M. (2012). Processing morphologically conditioned word accents. *The Mental Lexicon*, 7(1), 77-89.
- Soto-Faraco, S., Sebastián-Gallés, N., & Cutler, A. (2001). Segmental and suprasegmental mismatch in lexical access. *Journal of Memory and Language*, 45(3), 412-432.
- Steinhauer, K., Alter, K., & Friederici, A. D. (1999). Brain potentials indicate immediate use of prosodic cues in natural speech processing. *Nature Neuroscience*, 2(2), 191-196.
- Tanenhaus, M. K., Spivey-Knowlton, M., Eberhard, K. M., & Sedivy, J. C., (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, 268, 1632-1634.
- Tanenhaus, M. K., Magnuson, J. S., Dahan, D., & Chambers, C. (2000). Eye movements and lexical access in spoken-language comprehension: Evaluating a linking hypothesis between fixations and linguistic processing. *Journal of Psycholinguistic Research*, 29(6), 557-580.

- Tremblay, A., Broersma, M., & Coughlin, C. E. (2017). The functional weight of a prosodic cue in the native language predicts the learning of speech segmentation in a second language. *Bilingualism: Language and Cognition*, 1-13.
- Ullman, M. (2005). A cognitive neuroscience perspective on second language acquisition: The declarative/procedural model. In C. Sanz (ed.), *Mind and context in adult second language acquisition: Methods, theory, and practice* (pp.141-178). Washington, DC: Georgetown University Press.
- Weber, A., Grice, M., & Matthew, C. W. (2006). The role of prosody in the interpretation of structural ambiguities: a study of anticipatory eye movements. *Cognition*, 99(2), B63-B72.
- Weber-Fox, C. & Neville, H. (1996). Maturational constraints on functional specialization for language processing: ERP and behavioral evidence in bilingual speakers. *Journal of Cognitive Neuroscience*, 8, 231-256.
- Wechsler, D. (1997). Wechsler Adult Intelligence Scale (3rd ed.). San Antonio, TX: Psychological Corporation.

[illegible]

Table 2. One-sided t-tests of gating data for response accuracy to oxytone and paroxytone targets.

[illegible]

Table 3. Model output for gating data. Accuracy as a function of group and working memory.

Effect	β	Std. Error	CI lower	CI upper	z-value	Sig.
Intercept	1.44	0.15	1.14	1.74	9.36	< 0.001
LA	0.03	0.20	-0.36	0.42	0.14	= 0.88
LB	-1.06	0.24	-1.53	-0.58	-4.36	< 0.001
Working memory	0.14	0.08	-0.02	0.30	1.69	= 0.09

Fig. 1. Sample trial in the eye-tracking task.

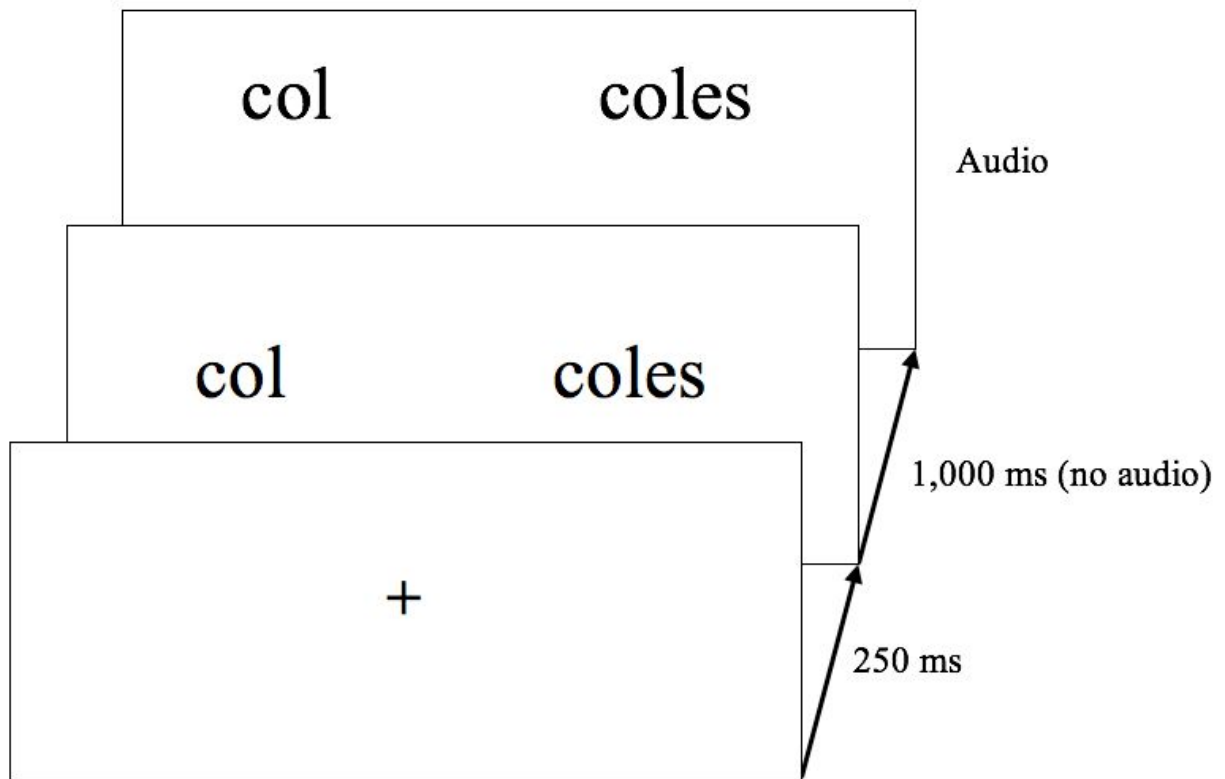


Figure 2. Output for final mixed effects model of eye tracking data. Model tested proportion of target fixations as a function of group and syllable structure. Native Spanish speakers represent the baseline (intercept).

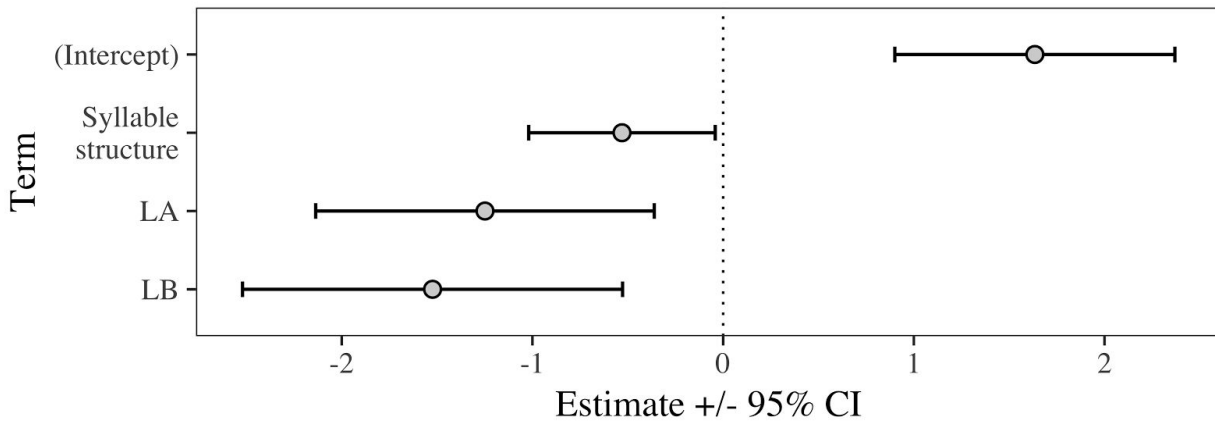


Figure 3. Proportion of target fixations from 750 ms before target syllable offset to 550 ms after as a function of group and syllable structure. Symbols and point ranges represent means \pm SE. The dotted vertical lines represent the mean target word onset and the target syllable offset.

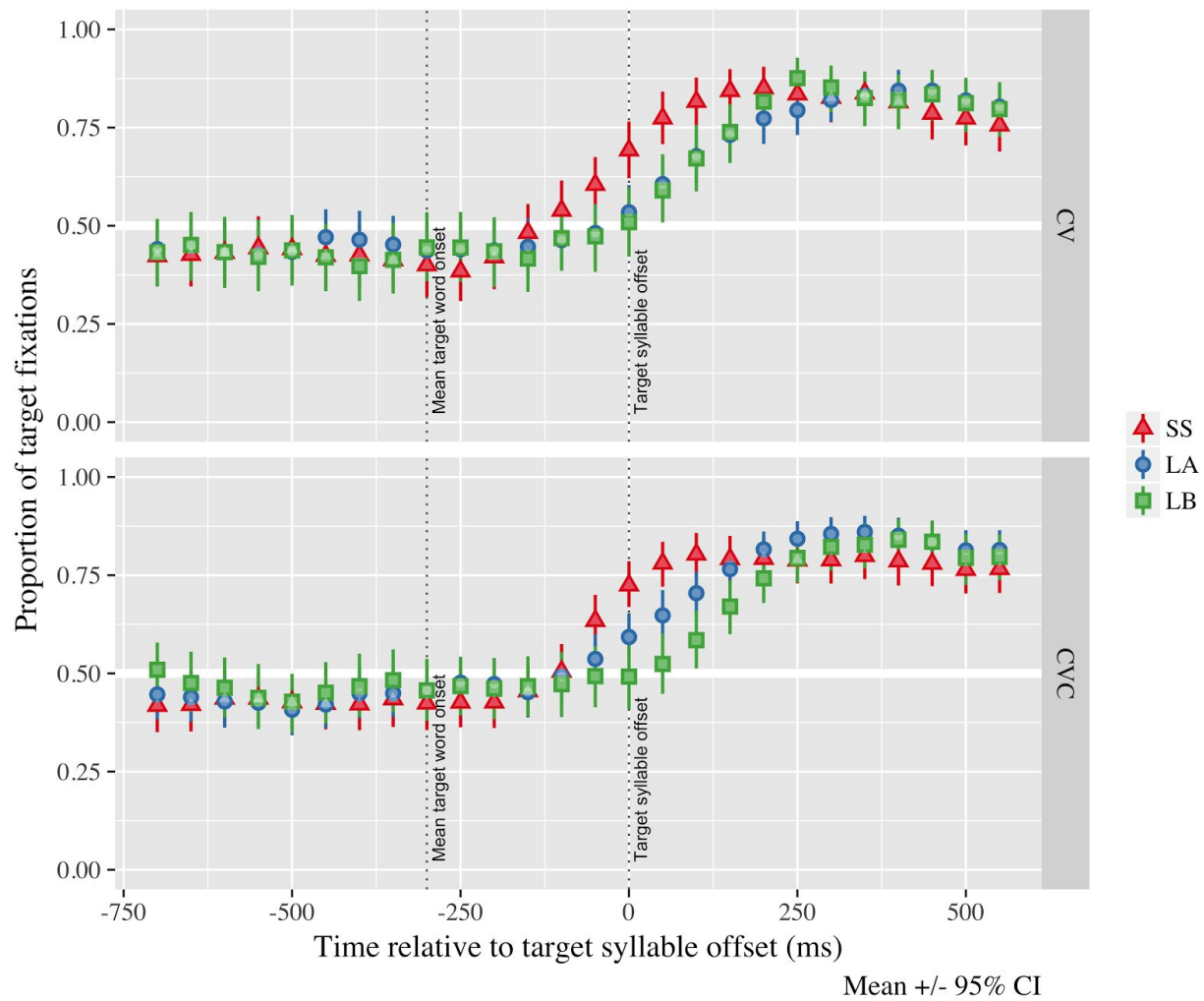


Figure 4. Observed target fixation empirical log-odds and growth curve model fits for effect of group and syllable structure during analysis window. Symbols and point ranges indicate mean \pm SE. Lines represent model fits.

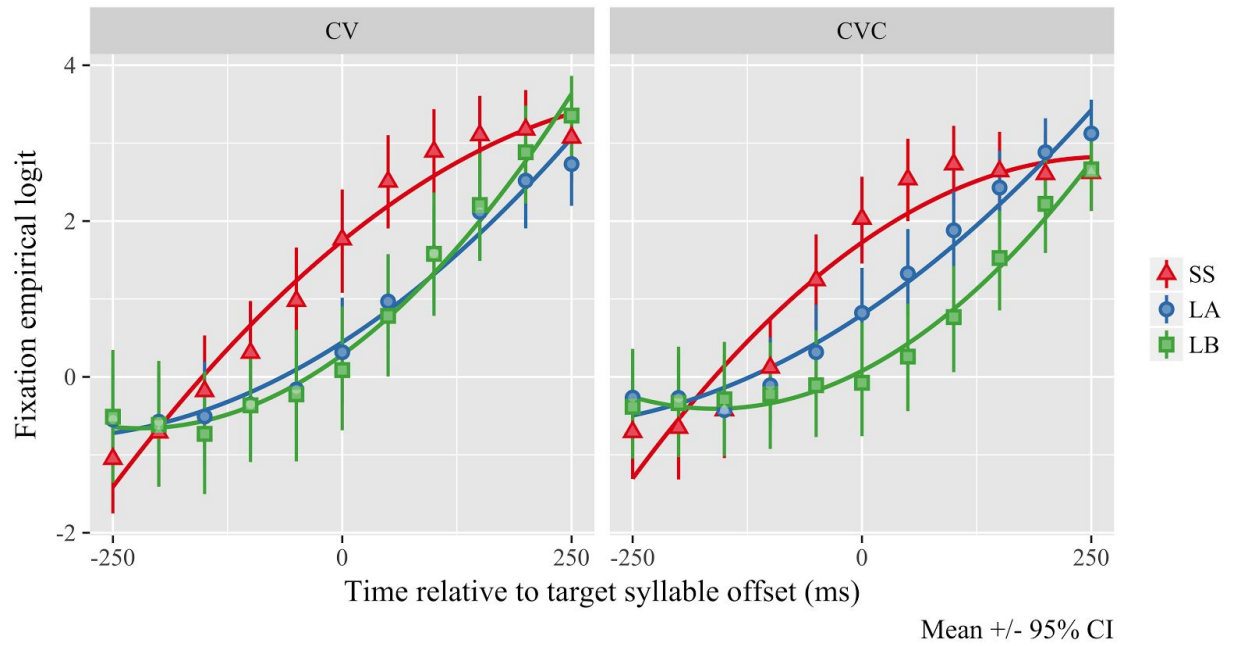
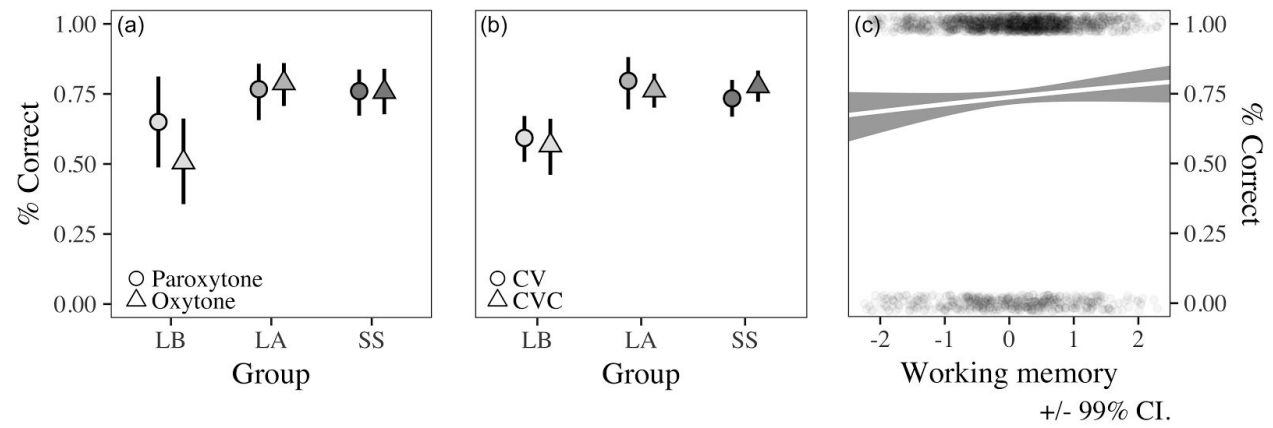


Figure 5. Proportion of correct responses on the gating task as a function of lexical stress (a), syllable structure (b) and working memory (c) for each group.



Appendix A

Final model output from growth curve analysis of eye-tracking data.

Random effects

Groups	Name	Variance	Standard Deviation	Corr	
Participant:Stress	ot1	3.53	1.88		
	ot2	0.77	0.88	0.18	
Participant:Syllable Struc.	(Intercept)	0.19	0.44		
	ot1	1.33	1.15	-0.35	
	ot2	0.09	0.29	-0.45	0.59
Participant	(Intercept)	0.25	0.49		
	ot1	0.25	0.50	0.85	
	ot2	0.08	0.29	0.80	0.37
Residual		9.04	3.00		

Fixed effects

Term	Estimate	Standard Error	df	t value	Pr(> t)	
(Intercept)	0.89	0.15	8.50E+01	5.91	6.94E-08	***
ot1	3.29	0.38	8.60E+01	8.73	1.80E-13	***
ot2	-0.20	0.23	6.80E+01	-0.88	0.38	
Group LA	-0.37	0.20	8.50E+01	-1.84	0.07	.
ot1:Group LA	-0.80	0.51	8.60E+01	-1.58	0.12	
ot2:Group LA	1.07	0.32	6.90E+01	3.40	0.00	**
Group LB	-0.57	0.22	8.60E+01	-2.59	0.01	*
ot1:Group LB	-1.02	0.56	8.70E+01	-1.82	0.07	.
ot2:Group LB	1.04	0.35	7.10E+01	3.02	0.00	**
Syllable structure	0.07	0.05	1.06E+04	1.42	0.16	
ot1:SS	0.25	0.18	9.83E+03	1.40	0.16	
ot2:SS	0.05	0.17	6.70E+03	0.27	0.79	
Group LA:SS	-0.18	0.07	1.07E+04	-2.56	0.01	*
ot1:Group LA:SS	-0.34	0.24	9.91E+03	-1.42	0.16	
ot2:Group LA:SS	0.10	0.23	6.82E+03	0.44	0.66	
Group LB:SS	-0.02	0.08	1.06E+04	-0.20	0.84	
ot1:Group LB:SS	0.14	0.27	9.61E+03	0.52	0.60	
ot2:Group LB:SS	0.06	0.26	7.00E+03	0.23	0.82	