

# Suprasegmental information cues morphological anticipation during L1/L2 lexical access

Nuria Sagarra and Joseph V. Casillas  
Rutgers University

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 did not use prosodic information to anticipate word suffixes in any task or condition. 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 gain anticipatory processing patterns qualitatively, but not quantitatively, similar to monolinguals.

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

Anticipation is the pre-activation of upcoming information, and its role in L1 sentence processing has been known for decades (see Kamide, 2008; Huettig, Rommers, & Meyer, 2011, for overviews). In contrast, L2 anticipatory studies are recent and yield mixed findings. Thus, some studies show that morphological cues are used to anticipate (a) morphosyntactic information in beginning learners (Liburd, 2014) and advanced learners (Dussias, Kroff, Guzzardo Tamargo, & Gerfen, 2013 with Anglophones; Marull, 2017), as well as (b) semantic information

in intermediate learners (Foucart, Martin, Moreno, & Costa, 2014; Trenkic, Mirkovic, & Altmann, 2014) and advanced learners (Dijkgraaf, Hartsuiker, & Duyck, 2016; Peters, Grüter, & Borovsky, 2015). In contrast, other studies indicate that morphological cues are not integrated to predict (a) morphosyntactic information in intermediate learners (Dussias et al., 2013 with Italians; Lew-Williams & Fernald, 2010; Marull, 2017; Mitsugi & MacWhinney, 2016) and advanced learners (Grüter, Lew-Williams, & Fernald, 2012; Hopp, 2015), as well as (b) semantic information in advanced learners (Martin et al., 2013).

Importantly, most L2 anticipatory studies focus on morphological cues (explicitly taught) and ignore acoustic cues (implicitly learned). In particular, prosody (i.e., intonation, tone, stress, rhythm) is essential to process words (McQueen, 2005) and sentences (Cutler, Dahan, & van Donselaar, 1997). Yet, there are only a handful of L1 studies on prosodic cues and anticipation of inflectional morphology in upcoming words (see Hawthorne & Gerken, 2014, for a review), and within a word (Rehrig, 2017; Roll, Horne, & Lindgren, 2010; Roll, 2015; Söders-tröm, Horne, & Roll, 2017; Söderström, Roll, & Horne, 2012). Crucially, to our knowledge, there are only three L2 studies on prosodic cues and anticipation within a word, and they yield mixed results. Some reveal that adult learners can use prosodic cues to predict morphological information (advanced: Schremm, Söderström, Horne, & Roll, 2016), whereas others show they can't (beginners: Gosselke Berthelsen et al., 2018; high intermediates: Rehrig, 2017). None of these studies considers individual differences in L2 proficiency and WM. We fill this gap by investigating the role of proficiency and WM in how monolinguals and adult beginning and advanced learners integrate verb stem lexical stress cues to predict present-past suffixes.

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 processing reasons (see Sagarra & Herschensohn, 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

Anticipation is crucial to sentence comprehension because it permits rapid and incremental integration of incoming information (Altmann & Mirkovic, 2009), and promotes efficient communication (Jaeger & Snider, 2013). There is ample evidence that monolinguals make syntactic predictions based on prosodic cues, such as intonation (Steinhauer, Alter, & Friederici, 1999; Weber, Grice, & Matthew, 2006), tone (Roll, Horne, & Lindgren, 2011), vowel duration (Rehrig, 2017), and pauses (Hawthorne & Gerken, 2014; Kjelgaard & Speer, 1999). Other studies also show that monolinguals utilize fine-phonetic detail, such as co-articulatory differences in a determiner, to anticipate incoming words (Salverda, Kleinschmidt, & Tanenhaus, 2014). However, little is known about monolinguals' 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). 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

Most L2 anticipatory studies examine the role of L2 morphological cues on anticipating upcoming morphological information (e.g., S-V agreement, D-N-Adj agreement), an association explicitly and constantly covered by teachers. In contrast, learners must learn prosody-suffix connections through implicit mechanisms. The teachers' negligence of the prosody-suffix connection extends to research scholars. Thus, only three studies have explored the role of prosodic cues to predict upcoming suffixes, and they have produced opposite findings: Schremm et al. (2016) reported that L2 learners can use prosodic cues to anticipate suffixes, whereas Gosselke Berthelsen et al. (2018) and Rehrig (2017) found that they cannot.

Schremm et al., (2016) asked adult relatively proficient non-tone L1 learners of Swedish and Swedish monolinguals to listen to sentences in which the tone on the verb stem matched or mismatched the present or past tense suffix. Mismatched tone-suffix trials produced longer RTs in a verb identification task for both groups, suggesting that all the participants used tone cues to predict verb inflections. Matched tone-suffix trials generated a bigger advantage for monolinguals

than for learners, and within the L2 group the advantage increased with higher language experience.

In an follow-up study, Gosselke Berthelsen et al. (2018) examined whether beginning German-speaking learners of Swedish used Swedish word tones to predict morphological information within a noun. They found that they did not use the L2 prosodic cues, but that they had mastered several precursory steps that could potentially lead to sensitivity to the L2 cue at higher proficiency levels. They concluded that beginning L2 learners are not able to use cues involving prosody, semantic and morphology.

In a related study, Rehrih (2017) investigated whether high intermediate Chinese-English learners and English monolinguals used vowel duration differences in verb stems that were one-syllable (*pushed*) (longer *u*) or two-syllable long (*pushing*) (shorter *u*) to predict the suffix. The participants 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 monolinguals used vowel duration information to anticipate suffixes, but that the learners did not. An explanation for the null finding may be that the learners were not proficient enough (proficiency was measured via subjective self-ratings instead of an objective test), or due to the comparison of two syntactic constructions known to be learned early (active) and late (passive). We contribute to the findings of these two studies by examining another prosodic cue (lexical stress), language (Spanish), and multiple proficiency levels (beginners, advanced) to truly test language exposure effects.

## WM and L2 anticipation

Studies exploring the function of individual differences in WM on linguistic anticipation are not only scant but contradictory and limited to monolingual populations. WM is the activation and maintenance of short-lived memory items while performing sometimes complex and time-consuming cognitive tasks (Barrouillet & Camos, 2007). Huettig and Janse (2016) examined WM and processing speed effects on Dutch monolinguals' use of morphological cues to predict morpho-syntactic information, and they found that enhanced WM and faster processing speed predicted anticipatory eye movements. Otten and Van Berkum (2009) also investigated WM effects on anticipation in Dutch monolinguals' integration of morphological cues to anticipate morphosyntax, but they used ERPs and reported no WM effects. We extend these studies 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

Lexical stress (henceforth stress) is the relative prominence of a syllable with respect to other syllables in a word, and is associated with various acoustic correlates, mainly  $F_0$  (Hz), duration (ms) and intensity (dB) (see Gordon & Roettger, 2017, for a comprehensive review). In Spanish and English, stress is variable and phonologically contrastive (e.g., Spanish: *PApa* ‘potato’ vs. *paPA* ‘father,’ English: *PREsent* vs. *preSENT*), but Spanish is syllable-timed (syllables have a similar duration), whereas English is stress-timed (intervals between stressed syllables have similar duration) (see Hualde, 2005, for a review). 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, in spite of the fact that it is contrastive in both languages. Two cross-modal priming studies suggest that these difficulties might be explained by the fact that stress is processed differently in Spanish (Soto-Faraco, Sebastián-Gallés, & Cutler, 2001) and English (Cooper, Cutler, & Wales, 2002). In these studies, Spanish and English monolinguals listened to a prosodically matched (Spanish: *prinCI-prinCIpio*; English: *ADmi-ADMiral*) or mismatched prime (Spanish: *PRINci-prinCIpio*; English: *admi-Admiral*), and pressed a yes-no button as soon as they could determine whether the sound file corresponded to the word in the screen (Spanish: *PRINcipe-prinCIpio* ‘prince-beginning,’ English: *ADMiral-adMiration*). The results indicated that both groups showed match facilitation (faster RTs in matched than control conditions), but only the Spanish monolinguals showed mismatch inhibition (slower RTs in mismatched than control conditions). The authors suggest that this finding is due to the difference in the functional load of stress in English (weak, i.e., there are few stress minimal pairs that are semantically unrelated) versus Spanish (strong) (See Cutler, 2012). Thus, during lexical access Spanish listeners have to attend to stress to reduce competition, but English listeners do not. Unstressed vowel reduction is sufficient for lexical differentiation (Cutler, 2012; Tremblay, Bruesma, & Couhlin, 2017).

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 can be weighted differently, and that 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 (Bullock & Lord, 2003) and perceiving (Ortega et al., 2013; Saalfeld, 2012) 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 not use prosodic information to predict suffixes, 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., Sagarra & Herschensohn, 2010; Osterhout et al., 2008; Rossi et al., 2006), and on the use of morphological cues (Dussias et al., 2013; Marull, 2017) and prosodic cues (Schremm et al., 2016) to predict morphological information. 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 (eye-tracking task) and accuracy (eye-tracking and gating tasks), provided that they use stress cues to predict morphosyntactic information. We base this predictions on studies showing WM effects on L1 anticipatory eye movements (Huetting & Jansen, 2016), and L2 morphosyntactic processing (Sagarra, 2008; Sagarra & Herschensohn, 2010; Faretta-Stutenberg, 2014; Havik et al., 2009; Keating, 2010). It is noteworthy that our hypotheses relate to WM and anticipatory abilities, not to WM and overall lexical processing.

## 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 (over 56); however, the distribution was bimodal. Thus we divided the learners into two groups: (1) late beginning learners (LB), with scores of  $16.10 \pm 3.95$  standard deviations (*SD*) and (2) late advanced learners (LA), with scores of  $46.27 \pm 4.09$  *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 commu-

nity 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).

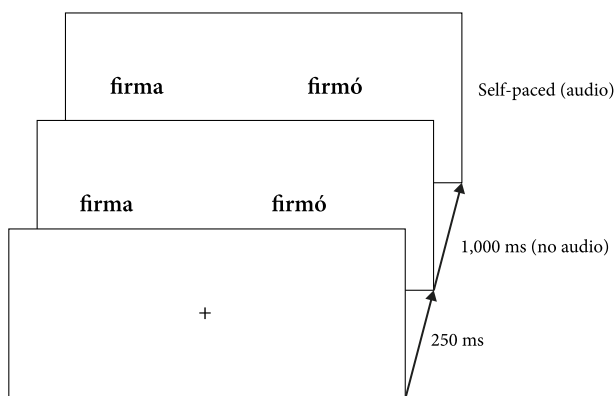


Figure 1. Sample trial in the eye-tracking task

They were instructed to press a 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, and transitive. 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  $-18\text{dB}$ , 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  $3.03 \pm 0.49$  SD syllables per second. The average length of the sentences was  $2.51 \pm 0.22$  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

This task 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, Magnuson, & Tahenhaus, 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  $\mu$  at 50% and the alternative hypothesis was that target fixations were greater than  $\mu$ . 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).

**Table 1.** One-sided t-tests eye-tracking data for proportion of target fixations at the offset of oxytone and paroxytone target syllables

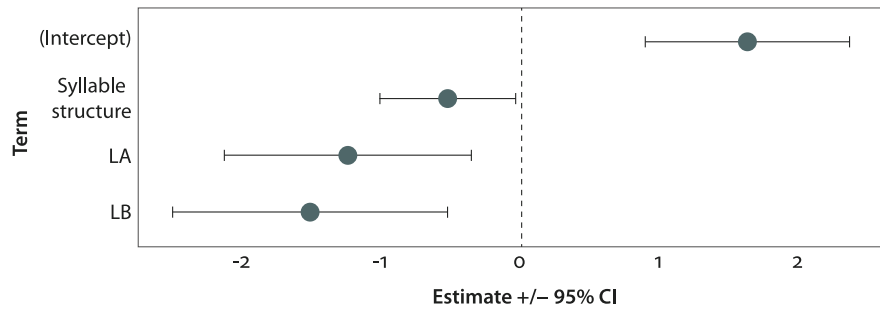
Group	Stress	Estimate	Statistic	CI low	Parameter	Sig.
SS	Paroxytone	0.63	2.80	0.55	21	< 0.005 *
	Oxytone	0.72	4.77	0.64	21	< 0.001 *
LA	Paroxytone	0.50	0.05	0.42	26	> 0.008
	Oxytone	0.58	2.13	0.52	26	= 0.02
LB	Paroxytone	0.50	0.02	0.40	18	> 0.008
	Oxytone	0.49	-0.21	0.39	18	> 0.008
Alpha = 0.008						

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 generalized linear mixed effects models (GLMM) 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  $1.25 \pm 0.34$  se ( $z = -3.62$ ,  $p < 0.001$ ), and for the LB group the log odds were reduced by  $1.52 \pm 0.39$  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 2 illustrates the final model output.



**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).

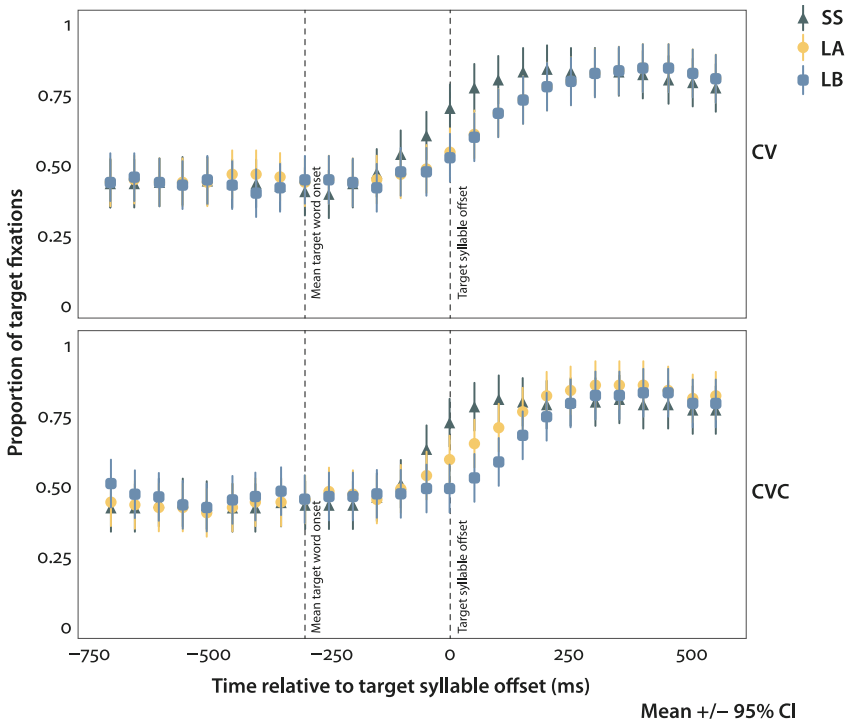
Finally, we used Growth Curve Analysis (Mirman, 2014) to analyze the time course of fixation from 500 ms before the offset of the initial syllable of the target word to 250 ms after (i.e., from the average target word onset to when the target fixations plateaued). We applied the empirical logit transformation (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:

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

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 devia-

tion 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 3. 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.

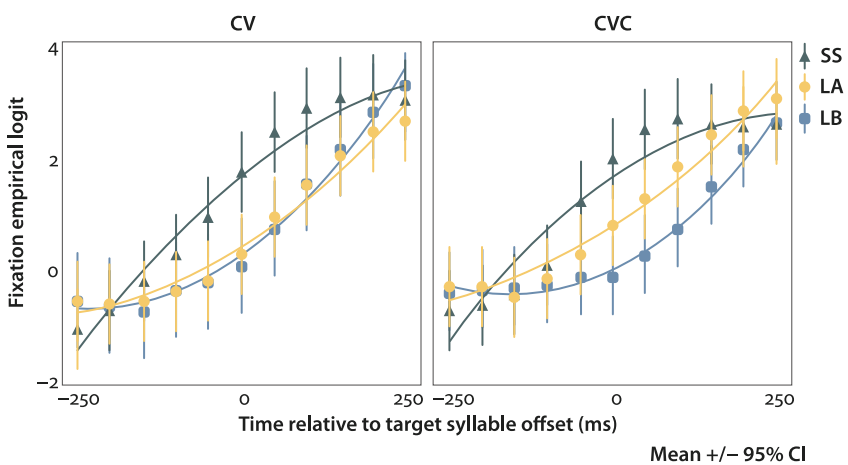


**Figure 3.** Proportion of target fixations from 750 ms before target syllable offset to 550 ms after as a function of group and syllable structure.

Symbols and point ranges represent means  $\pm$  SE. The dotted vertical lines represent the mean target word onset and the target syllable offset.

Model estimates are plotted in Figure 4. 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 3 and 4, 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 B.



**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.

### 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  $\mu$  at 50% accuracy, and the alternative hypothesis was that scores were greater than  $\mu$ . 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.

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

Group	Stress	Estimate	Statistic	CI low	Parameter	Sig.
SS	Paroxytone	0.76	7.57	0.70	24	< 0.001 *
	Oxytone	0.75	7.02	0.69	24	< 0.001 *
LA	Paroxytone	0.77	6.71	0.70	25	< 0.001 *
	Oxytone	0.79	8.87	0.73	25	< 0.001 *
LB	Paroxytone	0.65	2.15	0.52	9	= 0.03
	Oxytone	0.51	0.09	0.38	9	= 0.47
Alpha = 0.008						

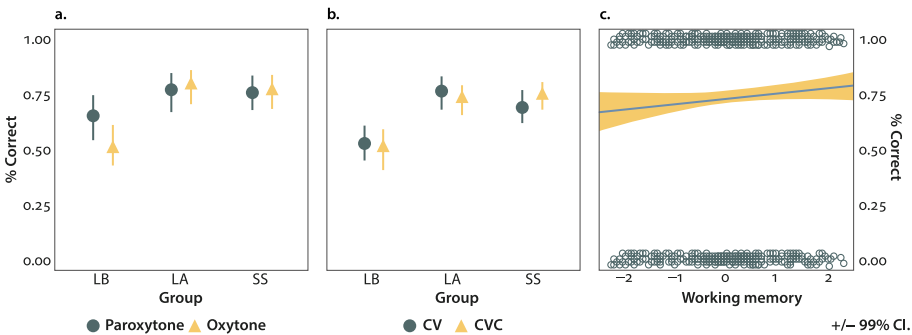
The objective of the second analysis of the gating data was to determine if response accuracy varied as a function of the fixed effects group (SS, LA, LB), stress (oxytone, paroxytone), syllable structure (CV, CVC) and standardized WM scores. Given the categorical nature of the outcome variable (correct, incorrect), the data were analyzed using 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 ( $R^2m = 0.03$ ,  $R^2c = 0.27$ ). 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  $1.06 \pm 0.24$  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.

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

Effect	Estimate	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



**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.

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). 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, and that they use suprasegmental cues to anticipate suffixes. Furthermore, this prosody-morphology interaction is independent of lexical meaning and is so strong that prosody can pre-activate morphology even when morphology is absent (i.e., replaced by a cough).

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 used differently in the L1 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's first syllable had a CVC structure providing 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. It is not clear if the relevant acoustic information corresponds with the coda segment, or with the increased duration of the target syllable. 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 that allowed the participants 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 (Dussias et al., 2013; Marull, 2017), but reduced or no integration at low proficiency levels (Dussias et al., 2013; Hopp, 2015; Lew-Williams & Fernald, 2010; Mitsugi & MacWhinney, 2016; Marull, 2017). Most importantly, our findings are in line with studies reporting that advanced (Schremm et al., 2017), but not beginning learners (Gosselke Berthelsen et al., 2018) integrate suprasegmental cues to anticipate morphological information within a word.

In addition to informing lexical access models by showing that adults can integrate L2 suprasegmental cues used differently in the L1 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., Sagarra & Herschensohn, 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 used differently in their L1 and L2 to make predictions in their L2. Based on our findings, we argue that adult learners can gain anticipatory patterns that are qualitatively similar to monolinguals (both use stress to anticipate suffixes), but that 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, but the gating task being offline. Therefore, in an offline task, when given time and ample acoustic information, the advanced learners perform similarly to natives. In contrast, in an online task, we can see that the extra acoustic information provided by the CVC syllable structure speeds processing. In its absence (i.e., in CV syllable structures), 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 (*FIR*ma '(s)he signs') and without (*fir*MÓ '(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 due 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 (Huettig, 2015).

## Acknowledgements

We would like to thank Joan Borràs for extracting the WM and gating data, and Cristina Lozano and Nicole Rodríguez for their assistance in collecting the L2 data. We also thank the participants, as well as the audience at the 2017 International Symposium on Bilingualism and the 2017 Hispanic Linguistic Symposium for helpful comments.

## 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  
<https://doi.org/10.1006/jmla.1997.2558>
- Altmann, G. T., & Mirkovic, J. (2009). Incrementality and prediction in human sentence processing. *Cognitive Sciences*, 33, 583–609. <https://doi.org/10.1111/j.1551-6709.2009.01022.x>
- Barr, D. J. (2008). Analyzing “visual world” eyetracking data using multilevel logistic regression. *Journal of Memory and Language*, 59(4), 457–474.  
<https://doi.org/10.1016/j.jml.2007.09.002>
- Bartoń, K. (2016). MuMin: Multi-Model Inference. R package version 1.15.6.  
<https://CRAN.R-project.org/package=MuMin>
- 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.  
<https://doi.org/10.1093/acprof:oso/9780198570394.003.0004>
- Boersma, P. & Weenink, D. (2017). *Praat: Doing phonetics by computer* [Computer program]. Version 6.0.33. Retrieved from: <http://www.praat.org/> (26 September, 2017).

- Bullock, B., & Lord, G. (2003). Analogy as a learning tool in second language acquisition: The case of Spanish stress. In A. T. Pérez-Leroux & Y. Roberge (Eds.), *Romance linguistics. Theory and acquisition* (pp. 281–297). Amsterdam: John Benjamins.  
<https://doi.org/10.1075/cilt.244.21bul>
- 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. [https://doi.org/10.1044/2014\\_JSLHR-L-13-0279](https://doi.org/10.1044/2014_JSLHR-L-13-0279)
- Clahsen, H. & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics*, 27, 3–42. <https://doi.org/10.1017/S0142716406060024>
- 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. <https://doi.org/10.1177/00238309020450030101>
- Cutler, A. (2012). *Native listening: Language experience and the recognition of spoken words*. Cambridge, MA: The 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.  
<https://doi.org/10.1177/002383099704000203>
- Dijkgraaf, A., Hartsuiker, R. J., & Duyck, W. (2016). Predicting upcoming information in native-language and non-native-language auditory word recognition. *Bilingualism: Language and Cognition*, 20(5), 1–14.
- 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. <https://doi.org/10.1017/S0272263112000915>
- 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: Cascadia 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), 1237–1267.  
<https://doi.org/10.1515/LING.2006.040>
- 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, NY: Springer. [https://doi.org/10.1007/978-1-4612-2852-3\\_3](https://doi.org/10.1007/978-1-4612-2852-3_3)
- 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. <https://doi.org/10.1017/S1366728915000486>
- Gordon, M., & Roettger, T. (2017). Acoustic correlates of word stress: A cross-linguistic survey. *Linguistics Vanguard*, 3(1). <https://doi.org/10.1515/lingvan-2017-0007>
- Gosselke Berthelsen, S., Horne, M., Brännström, K. J., Shtyrov, Y., & Roll, M. (2018). Neural processing of morphosyntactic tonal cues in second-language learners. *Journal of Neurolinguistics*, 45, 60–78.

- Grosjean, F. (1980). Spoken word recognition processes and the gating paradigm. *Perception and Psychophysics*, 28(4), 267–283. <https://doi.org/10.3758/BF03204386>
- 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. <https://doi.org/10.1177/0267658312437990>
- Hawthorne, K., & Gerken, L. A. (2014). From pauses to clauses: Prosody facilitates learning of syntactic constituency. *Cognition*, 133(2), 420–428. <https://doi.org/10.1016/j.cognition.2014.07.013>
- 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. <https://doi.org/10.1121/1.2188377>
- Hopp, H. (2015). Semantics and morphosyntax in predictive L2 sentence processing. *IRAL*, 53(3), 277–306. <https://doi.org/10.1515/iral-2015-0014>
- Hualde, J. I. (2005). *The sounds of Spanish*. New York, NY: Cambridge University Press.
- Huetting, F. (2015). Four central questions about prediction in language processing. *Brain Research*, 1626, 118–135. <https://doi.org/10.1016/j.brainres.2015.02.014>
- Huetting, F., & 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. <https://doi.org/10.1080/23273798.2015.1047459>
- 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. <https://doi.org/10.1016/j.jml.2007.02.001>
- 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. <https://doi.org/10.1016/j.actpsy.2010.11.003>
- Ingvallson, 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. <https://doi.org/10.1017/S1366728911000447>
- 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. [https://doi.org/10.1016/S0010-0277\(02\)00198-1](https://doi.org/10.1016/S0010-0277(02)00198-1)
- Ito, A. (2016). *Prediction during native and non-native language comprehension: The role of mediating factors*. Unpublished doctoral dissertation. The University of Edinburgh, UK.
- 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. <https://doi.org/10.1016/j.cognition.2012.10.013>
- Jensen, J. T. (1993). *English phonology*. Amsterdam: John Benjamins. <https://doi.org/10.1075/cilt.99>
- Jiang, N. (2004). Morphological insensitivity in second language processing. *Applied Psycholinguistics*, 25, 603–634. <https://doi.org/10.1017/S0142716404001298>
- Jiang, N. (2007). Selective integration of linguistic knowledge in adult second language learning. *Language Learning*, 57(1), 1–33. <https://doi.org/10.1111/j.1467-9922.2007.00397.x>



- Kamide, Y. (2008). Anticipatory processes in sentence processing. *Language and Linguistics Compass*, 2(4), 647–670. <https://doi.org/10.1111/j.1749-818X.2008.00072.x>
- 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. <https://doi.org/10.1006/jmla.1998.2620>
- Lenth, R. V. (2016). Least-Squares Means: The R Package lsmeans. *Journal of Statistical Software*, 69(1), 1–33. <https://doi.org/10.18637/jss.v069.i01>
- 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. <https://doi.org/10.1016/j.jml.2010.07.003>
- Lew-Williams, C. & Fernald, A. (2010). 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.), *BUCLD 41: Proceedings of the 41st Boston University Conference on Language Development* (pp. 466–480). Somerville, MA: Cascadia Press.
- Martin, C. D., Thierry, G., Kuipers, J.-R., Boutonnet, B., Foucart, A. & Costa, A. (2013). Bilinguals reading in their second language do not predict upcoming words as natives do. *Journal of Memory and Language*, 69, 574–588. <https://doi.org/10.1016/j.jml.2013.08.001>
- Matin, E., Shao, K. C. & Boff, K. R. (1993). Saccadic overhead: Information-processing time with and without saccades. *Perception and Psychophysics*, 53, 372–380. <https://doi.org/10.3758/BF03206780>
- 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. <https://doi.org/10.1016/j.jml.2006.06.006>
- 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: Lawrence Erlbaum Associates.
- Mirman, D. (2014). *Growth curve analysis and visualization using R*. Boca Raton, FL: Chapman & Hall/CRC.
- Mitsugi, S., & MacWhinney, B. (2016). The use of case marking for predictive processing in second language Japanese. *Bilingualism: Language and Cognition*, 19(1), 19–35.
- Nakagawa, S. & Schielzeth, H. (2013). A general and simple method for obtaining R<sup>2</sup> from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, 4, 133–142. <https://doi.org/10.1111/j.2041-210X.2012.00261.x>
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, 9, 97–113. [https://doi.org/10.1016/0028-3932\(71\)90067-4](https://doi.org/10.1016/0028-3932(71)90067-4)
- Ortega Llebaria, M., Gu, H., & Fan, J. (2013). English speakers' perception of Spanish lexical stress: Context-driven L2 stress perception. *Journal of Phonetics*, 41, 186–197. <https://doi.org/10.1016/j.wocn.2013.01.006>

- Otten, M. & Van Berkum, J. J. A. (2009). Does working memory capacity affect the ability to predict words in discourse? *Brain Research*, 1291, 92–101.
- Park, D., Welsh, R., Marschuetz, C., Gutches, 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. <https://doi.org/10.1162/089892903322598094>
- Peters, R. E., Grüter, T., & Borovsky, A. (2015). Anticipatory and locally coherent lexical activation varies as a function of language proficiency. In *Proceedings of CogSci2015: The Annual Meeting of the Cognitive Science Society*.
- R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <<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). <https://doi.org/10.1515/lingvan-2017-0006>
- 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. <https://doi.org/10.1017/S0332586515000189>
- Roll, M., Horne, M., & Lindgren, M. (2010). Word accents and morphology – ERPs of Swedish word processing. *Brain Research*, 1330, 114–123. <https://doi.org/10.1016/j.brainres.2010.03.020>
- 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. <https://doi.org/10.1016/j.jneuroling.2008.06.001>
- Roll, M., Söderström, P., & Horne, M. (2013). Word-stem tones cue suffixes in the brain. *Brain Research*, 1520, 116–120. <https://doi.org/10.1016/j.brainres.2013.05.013>
- Saalfeld, A. K. (2012). Teaching L2 spanish stress. *Foreign language annals*, 45(2), 283–303. <https://doi.org/10.1111/j.1944-9720.2012.01191.x>
- Sagarra, N. (2008). Working memory and L2 processing of redundant grammatical forms. In Z. Han (Ed.), *Understanding Second Language Process* (pp. 133–147). Clevedon, UK: Multilingual Matters.
- Sagarra, N. & Herschensohn, J. (2010). The role of proficiency and working memory in gender and number agreement processing in L1 and L2 Spanish. *Lingua*, 20, 2022–2039. <https://doi.org/10.1016/j.lingua.2010.02.004>
- Saslow, M. G. (1967). Latency for saccadic eye movement. *Journal of the Optical Society of America*, 57, 1030–1033. <https://doi.org/10.1364/JOSA.57.001030>
- Schremm, A., Söderström, P., Horne, M., & Roll, M. (2016). Implicit acquisition of tone-suffix connections in L2 learners of Swedish. *Mental Lexicon*, 11(1), 55–75. <https://doi.org/10.1075/ml.11.1.03sch>
- 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, P., Horne, M., & Roll, M. (2017). Stem tones pre-activate suffixes in the brain. *Journal of Psycholinguistic Research*, 46(2), 271–280. <https://doi.org/10.1007/s10936-016-9434-2>
- Söderström, P., Roll, M., & Horne, M. (2012). Processing morphologically conditioned word accents. *The Mental Lexicon*, 7(1), 77–89. <https://doi.org/10.1075/ml.7.1.04soe>

- 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. <https://doi.org/10.1006/jmla.2000.2783>
- 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. <https://doi.org/10.1038/5757>
- 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. <https://doi.org/10.1017/S136672891700030X>
- Trenkic, D., Mirkovic, J. & Altmann, G. 2014, 'Real-time grammar processing by native and non-native speakers: constructions unique to the second language.' *Bilingualism: Language and Cognition*, 17(2), pp. 237–257.
- 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. <https://doi.org/10.1016/j.cognition.2005.07.001>
- 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. <https://doi.org/10.1162/jocn.1996.8.3.231>
- Wechsler, D. (1997). *Wechsler adult intelligence scale* (3rd ed.). San Antonio, TX: Psychological Corporation.

## Appendix A.

**CVC first syllable structure:** *cambia-cambió* “(s)he changes-changed,” *canta-cantó* “(s)he sings-sang,” *compra-compró* “(s)he buy-bought,” *firma-firmó* “(s)he signs-signed,” *guarda-guardó* “(s)he keeps-kept,” *lanza-lanzó* “(s)he throws-threw,” *manda-mandó* “(s)he sends-sent,” *pinta-pintó* “(s)he paints-painted,” *rompe-rompió* “(s)he breaks-broke.”

**CV first syllable structure:** *bebe-bebió* “(s)he drinks-drank,” *come-comió* “(s)he eats-ate,” *gra a-grabó* “(s)he records-recorded,” *lava-lavó* “(s)he washes-washed,” *llena-llenó* “(s)he fills-fell in,” *saca-sacó* “(s)he takes-took (a picture),” *sube-subió* “(s)he goes-went up.”

Appendix B. 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		
Participant	ot2	0.09	0.29	−0.45	0.59	
	(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	

## **Address for correspondence**

Nuria Sagarra  
Rutgers University  
5177 Academic Building, 15 Seminary Place  
New Brunswick, NJ 08901  
USA  
[nuria.sagarra@rutgers.edu](mailto:nuria.sagarra@rutgers.edu)

## **Co-author information**

Joseph V. Casillas  
Rutgers University  
[joseph.casillas@rutgers.edu](mailto:joseph.casillas@rutgers.edu)