Anticipation of verbal suffixed based on lexical stress in Spanish by L1 Mandarin and L1 English speakers

Table of Contents

Laura Fernández Arroyo1

1 Rutgers University

Author note

Correspondence concerning this article should be addressed to Laura Fernández Arroyo, Rutgers University - Department of Spanish and Portuguese, 15 Seminary Place, New Brunswick, NJ 08904, USA . E-mail: [laura.fdeza@rutgers.edu](mailto:laura.fdeza@rutgers.edu)

Anticipation of verbal suffixed based on lexical stress in Spanish by L1 Mandarin and L1 English speakers

# 1. Introduction

Humans tend to anticipate information: when watching a movie, when driving, when playing sports, when listening to music. We also anticipate information in hearing speech, when reading, or when producing a linguistic message, especially when we are immersed in our first language (L1). While prediction is so pervasive in our lives, the nature of prediction mechanisms is unknown. Here I investigate how speakers transfer their L1 predictive mechanisms onto a L2. In particular, how acoustic and function knowledge is transferred to use suprasegmentals as cues to morphosyntactic outcomes. Understanding how morphosyntactic prediction works is crucial because it modulates anticipation at other linguistic levels (Guajardo & Wicha, 2014).  
To address this goal, I examine verbe tense anticipation based off lexical stress in adult Spanish monolinguals and English and Mandarin Chinese intermediate and advanced learners of Spanish. The conditioning factors I consider are language experience-related: L1 transfer and L2 proficiency, and cognitive (i.e., working memory, WM). In a usage-based framework, language experience is important because it shapes grammar constantly (Bybee, 2006) and so it shifts a speaker’s attention to and inhibition of input.  
L1 transfer in particular can affect the type of information that a speaker pays attention to by default because that information is important in the L1 and thus is encoded and activated in a regular basis. This inhibition effect is called blocking (Kamin, 1969) and prevents the speaker from encoding L2 cues associated to an outcome, because a previous cue transferred from the L1 is already encoded as a predictor of that outcome. Sometimes the L1 encoded cue can successfully lead to the correct outcome, but these cue-outcome mappings rarely fully overlap cross-linguistically. In cases of partial overlap, it is unclear whether the similarity or the discrepancy is stronger in impacting language processing. L2 proficiency affects the variety of linguistic contexts, words and structures that a learner has been exposed to, and the frequency with which the learner has encountered them. In theory, the larger the exposure, the easier to create a cue-outcome mapping. Language experience thus continuously influences a speaker’s ability to make predictions in any language they speak. Cognitive factors like WM may also be responsible for the inability to make L2 predictions. There are few studies, however, delving into how WM capacities affect speakers’ abilities especially in L2s to generate the predictions.

# 2. Background

## 2.1. Language anticipation

In a predictive processing view of language, language processing is facilitated and preceded by prediction of linguistic information based on cues already available at different linguistic and non-linguistic levels (e.g., Federmeier, 2007; Wicha, Bates, Moreno, & Kutas, 2003). These cues are associated with an outcome and the mapping helps generate the correct prediction (DeLong, Urbach, & Kutas, 2005; Dikker & Pylkkanen, 2011). Speakers used a wide range of phonological, morphosyntactic, semantic and contextual cues to anticipate linguistic outcomes in their L1. Morphosyntactic anticipation in particular is very important because it modulates anticipation at other linguistic layers, like semantic choices (Guajardo & Wicha, 2014).  
L2 processing is oftentimes difficult, and L2 anticipation is not a typical event. Studying the influence of L1 prediction mechanisms on L2 anticipation might explain why learning and understanding an L2 is so troublesome, and why attaining native-like proficiency in an L2 is rare. This influence might be mediated by cognitive capacities. Understanding and processing a L2 is more cognitively taxing than speaking a L1 because the L2 overtakes the pool of executive functions usually applied in L1 anticipation (Linck, Osthus, Koeth, & Bunting, 2014), especially at non-proficient levels. So when the cognitive executive resources are excessively loaded, they are undermined and this overtaxing may also hindered L2 predictive processing.

### 2.1.1. L1 processing and anticipation.

Language anticipation is a productive and dynamic cognitive mechanism in our L1. Speakers use a wide range of morphosyntactic (Grüter, Takeda, Rohde, & Schafer, 2016), syntactic (Linzen & Jaeger, 2016), semantic (Kamide, Altmann, & Haywood, 2003; Pozzan, Gleitman, & Trueswell, 2016) and phonological cues to generate predictions on semantic (Altmann & Kamide, 1999), morphosyntactic (Grüter et al., 2016; Lew-Williams & Fernald, 2010) and syntactic-discoursive outcomes (Grüter & Rohde, 2013). Among the phonological cues we use, we encounter coarticulation (Salverda, Kleinschmidt, & Tanenhaus, 2014), intonation (Nakamura, Arai, & Mazuka, 2012; Weber, Grice, & Crocker, 2006), lexical stress (Correia, Frota, Butler, & Vigário, 2013; Sagarra & Casillas, 2018), pauses between clauses (Hawthorne & Gerken, 2014; Kjelgaard & Speer, 1999), vowel duration (Rehrig, 2017), and tone (Roll & Horne, 2015; Roll, Horne, & Lindgren, 2011). Phonological cues are oftentimes associated with word suffixes (Roll, Horne, & Lindgren, 2010; Sagarra & Casillas, 2018; Soto-Faraco, Sebastián-Gallés, & Cutler, 2001).  
Phonological cues exert an influence on the predictions that are generated in both proximal and distal sentential contexts. Lexical stress in English is signaled through proximal cues like syllable’s F0, duration, and amplitude, but also through distal cues like meter patterns (Brown, Salverda, Dilley, & Tanenhaus, 2015), such that distal cues interact with the proximal ones during lexical segmentation and word recognition (Brown, Salverda, Dilley, & Tanenhaus, 2011). As a consequence, violating lexical stress expectations created by distal cues in terms of location and syllabic structure elicit anormal brain activity associated with a more costly processing (Domahs, Wiese, Bornkessel-Schlesewsky, & Schlesewsky, 2008). Phonological phenomena are closely tied to morphosyntactic suffixes, with relationships that vary across languages. In Central Swedish, a low tone cues the singular suffix in a noun (*fisken* ‘fish[SG]’), and a high tone cues the plural suffix (*fiskar* “fish[PL],” Roll et al., 2010; Roll, Söderström, & Horne, 2013; Söderström, Horne, & Roll, 2015). Likewise, low tone cues present tense ( *skrämmer* ‘I scare’), and a high tone cues past tense (*skrämde* “I scared,” Söderström, Roll, & Horne, 2012; Roll & Horne, 2015). Similarly, lexically stressed syllables in Spanish cue present tense (*CANta* ‘he sings’), while lexically unstressed syllables cue past tense (*canTÓ* “he sang,” Sagarra & Casillas, 2018). The presence or absenece of lexical stress at the beginning of a noun in Spanish can also signal what noun should be activated when several segmental competitors are possible (*PRINcipe* “prince” vs. *prinCIPIO* “beginning,” Soto-Faraco et al., 2001). Lastly, shorter vowel duration in verbs in English is associated with active voice (‘the girl was pushing the boy’), whereas longer vowel duration is associated with passive voice (“the girl was pushed by the boy;” Rehrig, 2017).  
Individual variability in cognitive capacities (i.e., WM) does not exert a great impact on linguistic anticipation in typical adult individuals, at least in simple morphological structures or in speech (Ye & Zhou, 2008). The few online studies so far on the connection between WM and language anticipation generally suggest that WM variability is not a factor in determining the efficacy and characteristics of morphosyntactic prediction generation in Dutch (Otten & Van Berkum, 2009), Spanish (Sagarra & Casillas, 2018) or English (Pakulak & Neville, 2010; Tanner & Van Hell, 2014). Although Huettig and Mani (2016) did find in a different study that WM could account for variability in Dutch gender morphosyntax anticipation. Since this last study shares with the previous one either language studied, or structure researched, or methodology employed, none of these can explain the different results on the influence of WM on language anticipation.  
In sum, L1 speakers generate predictions at different linguistic levels. If, when and how WM could affect the capacity to generate L1 predictions has still not been established, although findings suggests WM plays a minor role. Behavioral performance and brain activity may largely be conditioned by different levels of linguistic complexity or linguistic experience, so that WM only affects linguistic anticipation in certain cases.

### 2.1.2. L2 processing and anticipation.

*L1 transfer and L2 proficiency*  
Anticipation in a L2 is not as straightforward as in a L1. Results are especially mixed in the morphosyntactic domain. Morphosyntactic anticipation has been researched with gender (Dussias, Kroff, Tamargo, & Gerfen, 2013; Hopp, 2016; Lew-Williams & Fernald, 2010), verbal tense (Sagarra & Casillas, 2018; Schremm, Söderström, Horne, & Roll, 2016), number (Marull, 2017), case (Hopp, 2015; Mitsugi & Macwhinney, 2016), and cross-linguistic form similarity. The mixed results have been accounted for by factors like L1 transfer, or L2 proficiency.  
Regarding gender, L2 German speakers may not be able to generate gender predictions (Hopp, 2016), while L2 Spanish speakers may be able to do so (Dussias et al., 2013), although only under specific linguistic circumstances (Lew-Williams & Fernald, 2010). Contrary to gender, number is more difficult to anticipate, so only advanced speakers can make number predictions (Spanish, Marull, 2017), and L2 case is not predicted (German, Hopp, 2015; Japanese, Mitsugi & Macwhinney, 2016). For verbal tense, only L2 learners at advanced levels of proficiency can generate predictions in Spanish (Sagarra & Casillas, 2018), but intermediate learners can already make predictions in Central Swedish (Schremm et al., 2016).  
Studies varied in whether morphosyntax was used as the cue, the outcome, or both. The L1s were also different, so L1 transfer could have interacted with the role of morphosyntax in determining whether the L2 speakers were able to generate gender morphosyntactic predictions. L2 proficiency and maybe even WM can also partially account for the ability to make L2 predictions. All these conditioning factors have often been entangled in research, so it is difficult to tease their inluence apart. In gender anticipation, having a similar morphosyntactic system in the L1 can help generate predictions in the L2, even at lower levels of proficiency. This is the case of L1 Italian speakers anticipating gender suffixes in L2 Spanish at beginner stages (Dussias et al., 2013). Extrapolating L1 knowledge is also the case of L1 English speakers using definiteness in articles to predict nouns at intermediate proficiency stages (Lew-Williams & Fernald, 2010). And L1 transfer can also be seen in L1 English speakers generating predictions in L2 Dutch more efficiently when they use cue determiners that share similar forms with their L1 determiners (Liburd, 2014). In contrast, lacking such L1 system may hinder the generation of L2 predictions. This would be the case of L1 English speakers generating gender suffix predictions in L2 Spanish (Dussias et al., 2013).  
The stability and nature of the encoded L2 morphosyntactic system may influence whether predictions are generated or not. In German, L2 speakers are able to generate comprehension gender predictions only when they are also able to use the L2 gender system target-like in production (Hopp, 2013, 2016). In L2 Spanish, speakers may only predict gender suffixes when the nouns are highly frequent (Lew-Williams & Fernald, 2010).  
Overcoming this anticipation hurdle might be facilitated when at least a phonological layer is involved and the L2 linguistic experience is larger. English is a poor language in terms of verbal morphology when compared to Spanish. However, when morphosyntax is cued phonologically through lexical stress, advanced L2 Spanish speakers overcome the obstacle and generate verb suffix predictions (Sagarra & Casillas, 2018). L2 speakers of Swedish are also able to generate tense suffix predictions based on tones beginning at intermediate stages of proficiency (Schremm et al., 2016).  
However, there are a few considerations to bear in mind when studying transfer effects in phonological cues, since having a similar structure in the L1 does not guarantee it can be used for anticipation in the L2. Whereas advanced L2 Spanish speakers whose L1 is English learn to use lexical stress as a cue (Sagarra & Casillas, 2018), L1 French speakers rarely if ever learn to encode lexical stress as an anticipatory cue in L2 Spanish, simply because they do not learn to discriminate it (Dupoux, Sebastián-Gallés, Navarrete, & Peperkamp, 2008) even if French has lexical stress. In contrast, L1 Cantonese and L1 Mandarin speakers learn to discriminate lexical stress in L2 English (Chen, 2013; Li, To, & Ng, 2017), and L1 Koreans also learn to a certain extent (Hualde & Kim, 2015; Lee, Shin, & Garcia, 2019), even though none of these three languages encodes lexical stress. Although it has not been investigated whether these three populations learn to use lexical stress as a cue, these results suggest that the suprasegmental structure in the L1 can not only affect positively or not affect processing of L2 suprasegmental structures, but also affect L2 processing negatively depending on the L1 suprasegmental characteristics. This hypothesis would align with previous results that suggest that L1 non-tonal speakers can learn to use L2 tones as suprasegmental cues to morphosyntax (Hed, Schremm, Horne, & Roll, 2019; Schremm et al., 2016). In the case that L2 speakers create the mappings between prosodic cue and morphosyntactic outcome thanks to L1 sound encoding transfer, the mapping may not always reach native-like efficiency (Perdomo & Kaan, 2019). This loss of efficiency has not been accounted for in any model of L2 phonological acquisition. There are some models nevertheless on L2 segmental acquisition that could be adapted to explain how L1 transfer affects L2 prosody-based linguistic anticipation. Here I focus on the Second Language Linguistic Perception model (L2LP, Escudero, 2005, 2009; Van Leussen & Escudero, 2015) because it was thought out to explain and predict L2 sound perception phenomena. In L2LP, L2 sounds are perceived through the L1 filter, that is, how they would be perceived if they were pronounced in the speaker’s L1. Therefore, the acoustic similarities and discrepancies between the two phonological systems will shape the development of the encoding in the bilingual’s mind. In L2 sound acquisition, a copy of the L1 system is created during the initial stages, this copy starts to adjust with exposure via the Gradual Learning Algorithm, a comparison of the L1 system and the perceived L2 sounds. The algorithm offers three possibilities: a new sound is assimilated to multiple L1 sounds, a new sound is perceived as similar to another one in the L1 system, or a new sound does not equal any category in the L1 so it requires a new category for itself. Recent revisions to the model have further proposed that perceiving is not the same thing as recognizing, but there is still not enough research in this direction to formulate any hypothesis. This model might be able to explain L2 suprasegmental acquisition and its use for predictive processing because it proposes that the ‘components’ of the L1 sound will affect how the L2 sound is perceived, so having practice perceiving a correlate in the L1 should in theory ease L2 perception of that same correlate. This hypothesis would be supported by L1 tonal speakers learning L2 tones more easily than L1 non-tonal speakers (Chan & Leung, 2019). The fact that L2 speakers can learn to perceive or discriminate L2 suprasegmentals like lexical stress or tones even when the suprasegmental is absent in their L1, does not necessarilty mean that the L2 speakers can further use the learned sound to generate anticipation. The recent revisions to the L2LP specifically offer a trampoline into accounting for this step from discrimination processing to anticipatory processing.  
A problem that this or other models of L2 phonological acquisition have is that they might not be able to answer how the function of the suprasegmental in the L1 affects acquisition of the L2 suprasegmental structure, and how the function could influence perception even if acoustically the L1 and L2 structures are different. Speakers of two languages that have a function in common might transfer the function also when the acoustic correlates that characterize the function in each language differ. If function transfer were the case, the models would need to be adapted to cover this transfer too, and to explain how function transfer and acoustic correlate transfer differ.  
*Working memory*  
Linguistic knowledge and experience affect language processing and anticipation, especially in L2s. Evidence for that is French L1 speakers struggling to process L2 lexical stress in Spanish (Dupoux et al., 2008), arguably because their French lexical stress encoding blocks acquisition of L2 lexical stress with different properties. However, even in these non- or barely sensitive individuals, there are individual differences. Those individuals displaying greater sensitivity to stress differences also retain stress information longer in their WM (Schwab, Giroud, Meyer, & Dellwo, 2020). We should therefore consider cognitive differences beside linguistic factors when studying L2 acquisition, processing and anticipation. WM is the cognitive skill that allows to store incoming information temporarily and process it so other complex cognitive actions can be executed (Baddeley, 2007). Three main models of WM stand out: Just and Carpenter (1992)’s domain-specific single-resource model, Baddeley (2007)’s domain-specific multiple-resource model, and domain-free connectionist models. In the domain-specific single-resource model (Just & Carpenter, 1992), WM capacity constrains language learning and processing. There is also a trade-off between the ability to process and to store linguistic information. This trade-off comes as a consequence of the competition amongst the different actions required for a shared pool of cognitive resources that need to be divided. Therefore, when a task depletes or overtax the WM capacity of a person, their storage capacity diminishes and their processing slows down. In domain-specific multiple-resource model (Baddeley, 2007’) WM capacity also constrains linguistic performance, but the cognitive resources are organized differently. A central executive controls three subsystems: the short-term storage phonological loop, the short-term storage visuospatial sketchpad, and the episodic buffer. These three slave systems have independent limited capacities. The two short-term memory systems focus on content, while the episodic buffer connects its sister systems with the long-term memory. The central executive coordinates the activity of the whole entity by filtering the information received, assembling information from different sources into meaningful episodes, regulating the flow of information among the subsystems, shifting tasks, and shifting retrieval strategies. Lastly, connectionist models are domain-free as language capacities are determined by domain-general cognitive abilities. Specifically, by the ability to select which information to pay attention to and which information to inhibit, regardless of the nature of that information. In connectionist models there is no difference between processing and storage. Instead, WM is the activation of part of the long-term memory according to short-term patterns of activations related to domain-specific stores (e.g., Cowan, 2016).

WM influences L2 processing but the specific nature of this influence has still not been completely figured out. Some studies have found facilitating effects of WM on morphosyntactic processing (e.g., Nuria Sagarra, 2016; Sanz, Lin, Lado, Stafford, & Bowden, 2016), while others have found no effect (e.g., Foote, 2011; Grey, Cox, Serafini, & Sanz, 2015). These differences might be explained by the interaction with factors explained above, like variability in L2 proficiency and the cognitive load imposed by the task and the WM test (Sagarra, 2017). Speakers at a lower proficiency might then be more susceptible to WM overload (Serafini & Sanz, 2016). In L2 anticipation, higher WM learners may anticipate more easily agreement elements than lower WM individuals (Reichle, Tremblay, & Coughlin, 2016). Although the communication context may condition this ability: WM may mediate morphosyntactic prediction only under cognitively taxing conditions (Lozano-Argüelles, 2020), although this finding is not definitive, as a previous study revealed no effecs of WM on L1 and L2 Spanish speakers’ capacity to anticipate verbal tense based on lexical stress (Sagarra & Casillas, 2018).  
The literature in L2 anticipation show that speakers can achieve some success in L2 anticipation depending on their L1, the cues they need to process, their level of proficiency, and maybe their WM capacities, but their performance will rarely be native-like (Perdomo & Kaan, 2019). A possible explanation for the varied results on L2 perception and anticipation might be found in what speakers are transferring from their L1 (acoustic knowledge vs. function knowledge) that interacts with the new L2 structures they need to encode. Whereas L2 speakers’ anticipation performance might depend on their ability to perceive the cues and what needs to be anticipated, asymmetries amongst studies and the lack of cognitive measures also difficult comparison of results. The lack of a common theoretical framework, the use of non-standardized measures to assess proficiency (self-assessment, Lew-Williams & Fernald, 2010), the variety of tasks (e.g. eye-tracking, Sagarra & Casillas, 2018; vs offline, Dupoux et al., 2008), a variety of L1s (Hed et al., 2019), and the unclear distinction of variables (e.g., Schremm et al., 2016) call for further research where the possible factors accounting for L2 anticipation patterns are better distinguished.

## 2.2. Linguistic phenomena

### 2.2.1. Lexical stress.

Stress is the prominence of a syllable that speakers hear relative to the other syllables in the prosodic word (Hualde, 2005). The characteristics of lexical stress, such as acoustic correlates or position within words, vary widely across languages. In English and Spanish, lexical stress has no fixed position, but it position changes from word to word playing a phonologically contrastive function at the lexical level. This function is much more typical in Spanish than it is in English. In English, lexical stress is used predominantly to distinguish heteronyms or pairs of verb-noun that have no segmental differences (to “proDUCE,” verb vs. “PROduce,” noun). In Spanish, lexical stress differentiates all kinds of word categories and information, such as verbal tense and person (*CANto* “I sing” vs. *canTÓ* “s/he sang”), or nouns (*PApa* “potato” vs. *paPÁ* “dad”), or nouns from verbs (*TÉRmino* “term” vs. *terMIno* “I finish” vs *termiNÓ* “s/he finished”).  
Different acoustic correlates, their importance or weight, and how they combine cause the acoustic realization of lexical stress in each specific language to vary. In Spanish, the most reliable cues to stress are pitch (F0), duration and intensity (Hualde, 2005; Ortega-Llebaria, 2006; Ortega-Llebaria & Prieto, 2007, 2009). Pitch is higher for stressed syllables and lower for unstressed syllables, and stressed syllables sound louder and are usually slightly longer. The main cues in English are vowel duration and quality (Cooper, Cutler, & Wales, 2002; Cutler, 1986). Thus, unstressed vowels are centralized.  
Thanks to the lexically contrastive function, lexical stress helps in activation of lexical entries in L1 Spanish (Soto-Faraco et al., 2001), such that a prosodically matching cue to the target ( *prinCI* > *prinCIpio*, “start”) results in shorter and more accurate activation times, when compared to mismatching cues ( *PRINci* > *prinCIpio* ‘start’). These results suggest that participants were using the suprasegmental cues in lexical stress to anticipate the lexical element. L1 English speakers, in contrast, may only be able to use lexical stress as a cue when more than one syllable of the word has already been heard (Cooper et al., 2002). Perdomo and Kaan (2019) provide evidence of activation through prediction in even earlier syllables in the L1 English speakers, however, more in line with the performance of the L1 Spanish speakers in Soto-Faraco et al. (2001)‘s study. This difference in performance among studies probably results from the cues that speakers use to discriminate lexical stress, and how these cues are weighted in the language. That is, English L1 speakers may be placing a larger reliance on duration to process Spanish lexical stress. This reliance would be transferred from L1 English processing. Spanish L1 speakers may be relying more on other cues, such as pitch and intensity, that are discarded by the English speakers because they are not as informative in English (Ortega-Llebaria, Gu, & Fan, 2013).  
The differences and similarities in cue weighting can influence lexical stress perception in an L2. For example, Cooper et al. (2002) found that the similar distribution of stress in Dutch and English helped L1 Dutch learners of English transfer their knowledge of lexical stress to process it properly in L2 English. But, in contrast, German speakers have more difficulty perceiving stress in another free lexically stressed language such as Spanish than L1 speakers (Schwab & Dellwo, 2016); again, maybe due to the acoustic correlates they use to discriminate the suprasegmental. And L1 English L2 Spanish speakers can only use lexical stress as a cue in Spanish at advanced levels of proficiency (Sagarra & Casillas, 2018). A possible explanation is that the setting of lexical stress in English did not match the Spanish lexical stress encoding and thus prevented them from using it properly until they reached a certain level of linguistic experience. Following this line of reasoning, Vickie and Andruski (2010) suggested that language background in the L1 can affect how lexical stress is perceived in an L2 and what correlates are used to discriminate the L2 stress.  
The studies above suggest that the acoustic properties of prosody are essential in processing and activating incoming linguistic information. The importance speakers assign to each cue might extend beyond perception and affect anticipation as well. Studies like Vickie and Andruski (2010)’ further suggest that speakers resort to their prosodic abilities in the L1 in order to process other prosodic structures in their L2 absent in their L1, such that Chinese speakers can use pitch knowledge to discriminate lexical stress in Spanish, even if they lack that suprasegmental in their L1. Extending this hypothesis, L2 speakers might be able to transfer acoustic knowledge and function and reassemble them into new prosodic structures that are encoded in the L2 lexicon and use it as cues for language anticipation. Specifically, L1 speakers of lexically free stressed languages may transfer the function knowledge, and L1 tonal speakers might be able to transfer pitch knowledge to encode L2 lexical stress that is based primarily on pitch. After lexical stress has been encoded in the L2 lexicon thanks to L1 pitch knowledge, L2 speakers might learn to use this new prosodic structure as the basis for L2 prediction so as to reduce the processing cognitive load.

### 2.2.2. Tone.

Tones are the pitch contour patterns of the voiced part in syllables (Chao, 1968). Many languages use tones, or changes in pitch-contour, at a phrasal level for pragmatic purposes. However, only a few use tones contrastively at a lexical level. The acoustic correlates for tones vary across languages: some use only pitch (e.g., most Mandarin Chinese dialects), whereas others also use length and/or register (e.g., Cantonese Chinese). Relevant to this with Mandarin Chinese speakers, in most Mandarin Chinese dialects (e.g., from Beijing and Tianjin), the main and oftentimes only acoustic correlate for tones is changes in pitch (F0) contour or changes in pitch height within a syllable (Gandour & Fromkin, 1978; Zhu & Wang, 2015). Mandarin tones facilitate word recognition (Malins and Joanisse (2010)), although segmental cues are the primary activator, especially vowels (Hu, Gao, Ma, & Yao, 2012; Sereno & Lee, 2015; Wiener & Turnbull, 2016). The role of tone is thus integration of the activated lexical item in the higher context.  
Knowledge of the nature and function of tones in the L1 can affect L2 tone learning both positively and negatively by providing a background knowledge to which learners resort to acquire the L2 tones and compare them. Li et al. (2017) examined the influence of the L1 tonal knowledge in the acquisition of L2 tones in children. These children were L1 Cantonese speakers learning L2 Mandarin, and they had issues in categorizing Mandarin tones 1 and 4, as these tones would be assimilated to the same tone 1 category in Cantonese. In the case of these children, being a native speaker to a tonal language helped them in the perception of Mandarin tones 2 and 3, but it hindered perception of other L2 tones because the knowledge transfer from L1 tones disagreed from the L2 tone structure and interfered with it.  
The role of tones in Mandarin is different from the role of lexical stress in Spanish. Mandarin speakers need to pay attention to the pitch variations in order to assign the correct tone to the word they are hearing. Since pitch variations are the basis for lexical stress in Spanish, Mandarin speakers might be able to transfer their sensitivity to pitch changes to process and use pitch to anticipate linguistic information more easily than English speakers. In comparison, pitch variation is not as important an acoustic correlate in English lexical stress, so L1 English speakers may not be sensitive to L2 pitch information in lexical stress, at least in beginning and intermediate stages of proficiency. Therefore, L1 English speakers need to learn to distinguish pitch variations since they can only transfer function knowledge, while Mandarin speakers may transfer pitch knowledge but need to learn function in using lexical stress to anticipate verbal tense in L2 Spanish.

# 3. This study

L2 speakers are able to transfer prosody encoding abilities from their L1 to their L2 (Krishnan, Gandour, & Bidelman, 2010), so they learn to discriminate and process L2 phonological information even when the phenomena is absent in their L1 (e.g., Li et al., 2017). Sometimes, they can even use acoustic and prosodic resources in the L2 to anticipate oncoming nouns (Perdomo & Kaan, 2019) and verbal morphosyntax (Lozano-Argüelles, Sagarra, & Casillas, 2019; Sagarra & Casillas, 2018; Schremm et al., 2016). Lozano-Argüelles et al. (2019) provided evidence that L1 English speakers could use lexical stress as an anticipation cue to verbal tense in L2 Spanish at advanced levels of proficiency. While English and Spanish share the presence of lexical stress, the acoustic correlates are different in each language, so L1 English speakers might presumably have transferred function knowledge. Speakers of languages without lexical stress may learn to perceive it too (Chen, 2013; Li et al., 2017), but in their case, it would be knowledge of the acoustic properties of their L1 that may help in the perception of the L2 phonological phenomenon (Li et al., 2017).  
Taking a step further, the current study aims to address the question of how the type of knowledge transferred affects language predictive processing. Specifically, this study first examines whether the acoustic properties of the L1 help not only in perceiving prosodic elements in a different language with different phonological structures but also in using those elements to anticipate oncoming linguistic information, even when the acoustic properties are used for different purposes in each language (i.e., for tone vs lexical stress). Second, this study also explores how acoustic transfer compares to function transfer, when the L1 shares the function of a suprasegmental structure, but its realization is different. The populations selected to address these questions were L1 Mandarin and L1 English speakers who speak L2 Spanish at a proficiency level ranging from intermediate to near native. They completed an eye-tracking task that measures their prediction abilities in Spanish when they use lexical stress as a cue to verbal tense suffixes. They were compared to a L1 Spanish control group. The research questions therefore are:  
*Do Spanish monolinguals, and intermediate and advanced Mandarin and English learners of Spanish use lexical stress to anticipate verbal suffixes in Spanish? If so, are prediction abilities mediated by verbal WM?*  
Our predictions are that monolingual Spanish speakers will anticipate verb tense when cued by stressed and unstressed syllables (Lozano-Argüelles et al., 2019; Sagarra & Casillas, 2018), but WM will not be responsible for individual variability (**???**; Otten & Van Berkum, 2009; Sagarra & Casillas, 2018). L2 speakers will generate the predictions at advanced levels of proficiency (Lozano-Argüelles et al., 2019; Sagarra & Casillas, 2018). L1 English speakers may acquire Spanish lexical stress through successful transfer of function knowledge and L1 Mandarin speakers through successful transfer of pitch knowledge (Li & Grigos, 2018), but none will provide any advantage over the other typo of transfer. Increased verbal WM is expected to affect extrapolation efficiency positively (Huettig & Mani, 2016) at an individual level, and to bring the start of the effects of cross-linguistic association forward in the L2 acquisition process.

# 4. Methods

## 4.1. Participants

30 monolingual speakers of Spanish (20 females), 42 L2 Spanish L1 English speakers (30 females), and 43 L2 Spanish L1 Mandarin speakers (33 females) participated in the study. All participants were living in Madrid, Spain at the time of data collection. English L1 speakers had been in a Spanish-speaking country for a mean of 31.5 months ( *SD* = 28.6) and the L1 Mandarin speakers for a mean of 43.9 months ( *SD* = 43.6). All speakers were raised in monolingual families in monolingual communities of their L1. The monolingual Spanish speakers were local to the Madrid area and spoke no other language fluently nor had lived abroad. The participants belonging to all L2 groups started learning Spanish after age 12. At the time of data collection, the L1 Mandarin speakers used Spanish a mean of 47.3% of the week ( *SD* = 21.2) and the L1 English speakers a mean of 34.4% ( *SD* = 16.6). The age range for all groups was 18-45 (Spanish: mean = 26.2, *SD* = 8.82, English: mean = 25.7, *SD* = 4.27; Mandarin: mean = 24.5, *SD* = 4.39). All participants had normal to corrected-to-normal hearing and vision, and no motor disability.

## 4.2. Materials

### 4.2.1. Screening tasks.

A proficiency test was administered to the L2 participants to screen them and to assign them a level of proficiency. The proficiency test was a shortened version (Sagarra & Herschensohn, 2010) of the *Diploma de Español como Lengua Extranjera* (‘Certificate of Spanish as a Foreign Language’, by Instituto Cervantes). The test consisted of 56 questions: 16 on grammar, 10 on vocabulary, and 20 on reading comprehension. Each question was worth one point. Participants who scored below 25 were disqualified for the study.  
Participants were interviewed and completed a language background questionnaire, to make sure that they qualified for the study. The background questionnaire contained questions related to their age, handedness, what language they spoke in their household while growing up, how old they were when they started learning Spanish, how long they had lived in Spanish-speaking countries, how much they used each language per week at the time of data collection, and other languages that they spoke fluently.  
Participants also completed a vocabulary test with some of the words from the linguistic anticipation task. This test was a multiple-choice test with 17 screens with 8 words each. For each word in Spanish, they had to choose between 8 possible meanings in the participants’ L1. An extra option was provided to mark in case they did not know the meaning of the word in Spanish.

### 4.2.2. Linguistic materials.

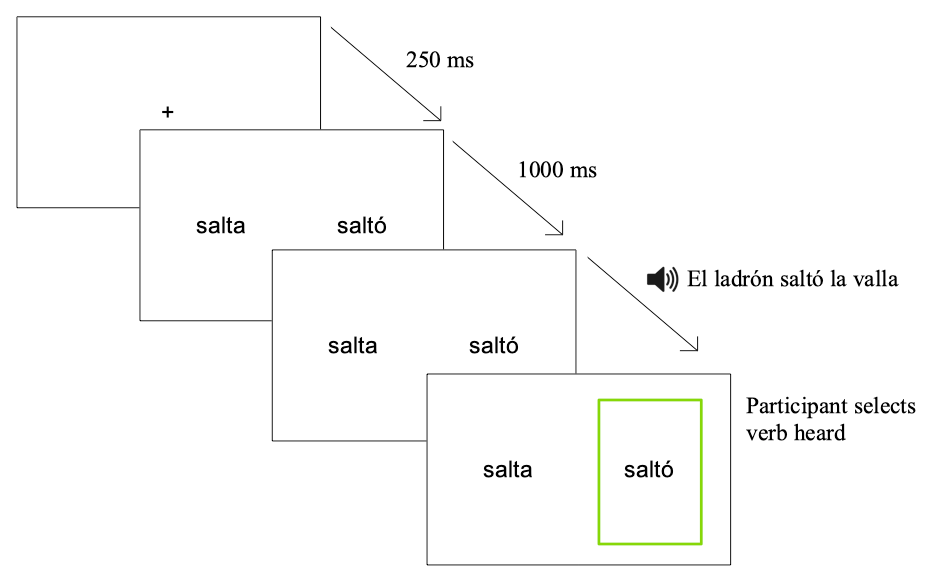
A visual-world paradigm was used to asses participants’ abilities to associate verbal suffixes with stressed or unstressed syllables. The eye-movements during the task were recorded through an EyeLink 1000 Plus desktop mount eye-tracker from SR Research (sampling rate: 1k Hz; spatial resolution was less than .05o; averaged calibration error: .25-.5o). Participants read two verbs on a computer screen side by side ( *salta* ‘s/he jumps,’ *saltó* ‘s/he jumped’) and heard a sentence containing one of the two verbs ( *El ladrón saltó la valla* ‘the thief jumped over the fence’). Their task consisted of selecting the verb they had heard as fast as possible by pressing the right- or left-shift key. When they made their selection, a green rectangle appeared around the selected verb. There were 4 practice sentences, 16 experimental sentences, and 80 fillers. The verbs were the present tense and preterit tense 3rd person singular forms of regular -ar verbs. Spanish regular verbs have 2 cues to tense: initial syllable stress and word ending morphology. Stressed initial syllables indicate that the tense is present ( *salta* ‘s/he jumps’), and unstressed initial syllables indicate that the tense is past (preterit, *saltó* ‘s/he jumped’). Regarding morphology, the verb suffix indicates tense, person and aspect. In total there were 16 verbs, so 16 pairs of present and preterit verbs, 16 in each condition, 32 final forms. All verbs had two syllables in both conditions: the first syllable had a CVC structure and the second syllable a CV structure. Participants never heard the two conditions for a single verb. The number of verbs in tense on each side of the screen was counterbalanced across trials and across participants.  
The aural sentences were 5 words long with a neutral structure NPSUBJ-V-NPOBJ. The sentences for each condition within a pair were the same. The sentences were distributed into blocks by means of a Latin square design. There were 8 blocks. Each block contained 2 experimental sentences, one of each type, and 6 filler sentences. The blocks appeared in a randomized order. Within and across the blocks, the sentences were pseudo-randomized to avoid two experimental sentences of the same condition appearing one after the other.

### 4.2.3. Verbal WM materials.

An aural version of Unsworth, Heitz, Schrock, and Engle (2005)’s Operation Span task (OSpan) was used to measure verbal WM capacities. This task generates measures of storage and processing speed. In a single trial, participants heard first a word and then a simple mathematical problem that could be either true (2 + 2 = 4) or false (2 + 2 = 5). Along the mathematical problem, they saw the words TRUE and FALSE on the screen. They had to select as fast as possible the correct word depending on whether what they heard was true or false by pressing the left- or right-shift key corresponding to the side on which their response was. This sequence of word-mathematical problem would repeat a certain number of times until a set was complete. At the end of each set, participants were prompted to write down the individual words they had heard before each problem in the same order they had heard them.  
There were two practice trials of 3 words and simple mathematical problems, and three experimental sets of three, four, five and six words and mathematical problems. The words and mathematical problems that participants heard in the sets appeared in a fixed order across participants. The words TRUE and FALSE that participants needed to select in response to the mathematical problems appeared counterbalanced on each side across participants. No feedback was provided. The whole task was administered in the participant’s L1.

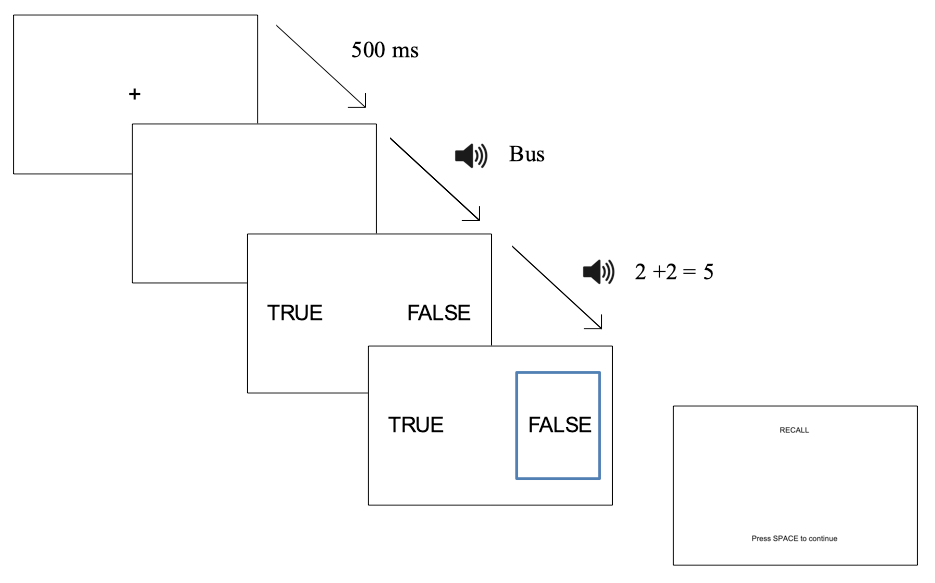
## 4.3. Procedure

Data collection was conducted in a single session with each participant individually. Each session lasted around 1 hour and 15 minutes and participants could take breaks between tasks anytime they wanted. Only the researcher accompanied the participant during this time and all interactions happened in Spanish. The room was isolated from external noise and light. Participants completed the tasks in this order: Spanish proficiency test (only L2 learners; 15-20 minutes), language background questionnaire (10 minutes), eye-tracking task (25 minutes), OSpan (15 minutes), and a vocabulary test (10 minutes). First, participants listened to an overview of the tasks and signed the consent form. Participants provided oral responses for the language background questionnaire. They completed the remaining tasks on a computer.  
*Visual-world paradigm.* For the visual-world paradigm, participants rested their head on a chin rest, completed an 11-point grid calibration task, and received task instructions both orally and in writing. Then, they completed the practice trials and were given the opportunity to ask questions. Afterwards, they performed the task. Both the practice and the task trials followed this order. First, participants looked at a fixation sign in the middle of the screen for 250 ms. This fixation sign allowed the researcher to recalibrate manually when necessary. Then, two words appeared in the screen side to side. Once the words had been on the screen for 1000 ms., the sentence started playing and continued until reaching the last word. That is, the sentence did not stop when participants selected the word they heard. Participants were instructed to select the word on the screen they heard on the sentence as fast as possible by pressing the right- or left-shift keys. A green rectangle appeared on the screen around the selected word when participants pressed the key to make their choice. Response recording was set up to be registered only when the press happened after the start of the verb in the sentence. Previous presses where not recorded so the setting forced participants to press again until they saw the green rectangle appear. No feedback was provided. See Figure 1 for an example of an experimental trial.



*Figure* *1:*. Sample trial of the visual-world paradigm

*OSpan* The OSpan task was divided into practice and experimental trials. For each trial, participants first heard a word and then heard a simple mathematical problem, all in their L1. During the equation, participants saw the words TRUE or FALSE on each side of the screen, and they had to press the corresponding key (left-shift key for the word on the left, right-shift key for the word on the right) depending on whether the problem was correct or not. After each problem, they saw a fixation point for 500 ms. and another word-equation pair was presented. This process was repeated until the word RECALL or a linguistic equivalent in the L1 of the participant was shown on the screen, at which moment participants had to write on a piece of paper the words they had heard for that set in the order they were presented. The task started with three sets of three trials, then three sets of four trials, and so on until the third set of six trials. Figure 6 shows a sample trial.



*Figure* *2:*. Sample trial of the OSpan

## 4.3. Statistical analysis

Statistical analyses were conducted on R (R Core Team, 2019) using the packages lme4 (Bates, Mächler, Bolker, & Walker, 2014) and multcomp (Hothorn et al., 2016). The gaze-fixation data was shifted 200 ms. to account for the time it takes to plan and launch a saccade (e.g., Fischer, 1992; Saslow, 1967), downsampled to 10 ms. and 50 ms. bins and centered at the onset of the last syllable of target items. The empirical logit transformation was applied to binary responses (fixations on target or distractor; Barr, 2008). The 10 ms. bins were run through independent t-tests to find out whether participants were anticipating the correct verbal tense upon hearing a stressed or unstressed initial syllable. The 50 ms. bins were used to model a growth curve analysis (Mirman, 2016) to observe how the pattern of gaze fixations changed as suprasegmental and segmental information became available. The models with the time course were implemented by using linear, quadratic, and cubic orthogonal polynomials with the independent variables group and lexical stress. The monolinguals served as baseline. Lexical stress had been sum-coded. By-subject and by-item random effects were also tested. Nested model comparisons were implemented to assess main effects and interactions. For the OSpan, reaction times to the mathematical problems were averaged and used as score, since different languages have differnt memory score and therefore are not comparable. The proficiency was analyzed as a continuous scale. OSpan scores were analyzed as homogeneity measure and as a continuous scales to observe how verbal WM capacity affects language processing and anticipation.

# References

Altmann, G. T., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, *73*(3), 247–264.

Baddeley, A. (2007). *Working memory, thought, and action* (Vol. 45). OUP Oxford.

Barr, D. J. (2008). Analyzing ‘visual world’eyetracking data using multilevel logistic regression. *Journal of Memory and Language*, *59*(4), 457–474.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv Preprint arXiv:1406.5823*.

Brown, M., Salverda, A. P., Dilley, L. C., & Tanenhaus, M. K. (2011). Expectations from preceding prosody influence segmentation in online sentence processing. *Psychonomic Bulletin & Review*, *18*(6), 1189–1196.

Brown, M., Salverda, A. P., Dilley, L. C., & Tanenhaus, M. K. (2015). Metrical expectations from preceding prosody influence perception of lexical stress. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(2), 306.

Bybee, J. (2006). From usage to grammar: The mind’s response to repetition. *Language*, *82*(4), 711–733. Retrieved from <http://www.jstor.org/stable/4490266>

Chan, R. K., & Leung, J. H. (2019). WHY are lexical tones difficult to learn?: INSIGHTS from the incidental learning of tone-segment connections. *Studies in Second Language Acquisition*, 1–27.

Chao, Y. R. (1968). A grammar of spoken chinese.

Chen, H. C. (2013). Chinese learners’ acquisition of english word stress and factors affecting stress assignment. *Linguistics and Education*, *24*(4), 545–555.

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.

Correia, S., Frota, S., Butler, J., & Vigário, M. (2013). Word stress perception in european portuguese. In *INTERSPEECH* (pp. 267–271).

Cowan, N. (2016). *Working memory capacity: Classic edition*. Psychology press.

Cutler, A. (1986). Forbear is a homophone: Lexical prosody does not constrain lexical access. *Language and Speech*, *29*(3), 201–220.

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience*, *8*(8), 1117–1121.

Dikker, S., & Pylkkanen, L. (2011). Before the n400: Effects of lexical–semantic violations in visual cortex. *Brain and Language*, *118*(1-2), 23–28.

Domahs, U., Wiese, R., Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2008). The processing of german word stress: Evidence for the prosodic hierarchy. *Phonology*, 1–36.

Dupoux, E., Sebastián-Gallés, N., Navarrete, E., & Peperkamp, S. (2008). Persistent stress “deafness”: The case of french learners of spanish. *Cognition*, *106*(2), 682–706.

Dussias, P. E., Kroff, J. R. V., Tamargo, R. E. G., & Gerfen, C. (2013). When gender and looking go hand in hand: Grammatical gender processing in l2 spanish. *Studies in Second Language Acquisition*, *35*(2), 353–387.

Escudero, P. (2005). *Linguistic perception and second language acquisition: Explaining the attainment of optimal phonological categorization*. Netherlands Graduate School of Linguistics.

Escudero, P. (2009). Linguistic perception of “similar” l2 sounds. *Phonology in Perception*, *15*, 152–190.

Federmeier, K. D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology*, *44*(4), 491–505.

Fischer, B. (1992). Saccadic reaction time: Implications for reading, dyslexia, and visual cognition. In *Eye movements and visual cognition* (pp. 31–45). Springer.

Foote, R. (2011). Integrated knowledge of agreement in early and late english–spanish bilinguals. *Applied Psycholinguistics*, *32*(1), 187–220.

Gandour, J. T., & Fromkin, V. (1978). The perception of tone in tone: A linguistic survey.

Grey, S., Cox, J. G., Serafini, E. J., & Sanz, C. (2015). The role of individual differences in the study abroad context: Cognitive capacity and language development during short-term intensive language exposure. *The Modern Language Journal*, *99*(1), 137–157.

Grüter, T., & Rohde, H. (2013). L2 processing is affected by rage: Evidence from reference resolution. In *The 12th conference on generative approaches to second language acquisition (gasla)*.

Grüter, T., Takeda, A., Rohde, H., & Schafer, A. J. (2016). L2 listeners show anticipatory looks to upcoming discourse referents. In *Poster presented at the 41st annual boston university conference on language development, boston university* (pp. 4–6).

Guajardo, L. F., & Wicha, N. Y. (2014). Morphosyntax can modulate the n400 component: Event related potentials to gender-marked post-nominal adjectives. *Neuroimage*, *91*, 262–272.

Hawthorne, K., & Gerken, L. (2014). From pauses to clauses: Prosody facilitates learning of syntactic constituency. *Cognition*, *133*(2), 420–428.

Hed, A., Schremm, A., Horne, M., & Roll, M. (2019). Neural correlates of second language acquisition of tone-grammar associations. *The Mental Lexicon*, *14*(1), 98–123.

Hopp, H. (2013). Grammatical gender in adult l2 acquisition: Relations between lexical and syntactic variability. *Second Language Research*, *29*(1), 33–56.

Hopp, H. (2015). Semantics and morphosyntax in predictive l2 sentence processing. *International Review of Applied Linguistics in Language Teaching*, *53*(3), 277–306.

Hopp, H. (2016). Learning (not) to predict: Grammatical gender processing in second language acquisition. *Second Language Research*, *32*(2), 277–307.

Hothorn, T., Bretz, F., Westfall, P., Heiberger, R. M., Schuetzenmeister, A., Scheibe, S., & Hothorn, M. T. (2016). Package “multcomp”. *Simultaneous Inference in General Parametric Models. Project for Statistical Computing, Vienna, Austria*.

Hu, J., Gao, S., Ma, W., & Yao, D. (2012). Dissociation of tone and vowel processing in m andarin idioms. *Psychophysiology*, *49*(9), 1179–1190.

Hualde, J. I. (2005). *The sounds of spanish with audio cd*. Cambridge University Press.

Hualde, J. I., & Kim, J. Y. (2015). The acquisition of spanish lexical stress by korean learners. In *ICPhS*.

Huettig, F., & Mani, N. (2016). Is prediction necessary to understand language? Probably not. *Language, Cognition and Neuroscience*, *31*(1), 19–31.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, *99*(1), 122.

Kamide, Y., Altmann, G. T., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language*, *49*(1), 133–156.

Kamin, L. (1969). Predictability, surprise, attention and cognition. *BA Campbell, RM Church New York: Appleton-Century-Crofts*.

Kjelgaard, M. M., & Speer, S. R. (1999). Prosodic facilitation and interference in the resolution of temporary syntactic closure ambiguity. *Journal of Memory and Language*, *40*(2), 153–194.

Krishnan, A., Gandour, J. T., & Bidelman, G. M. (2010). The effects of tone language experience on pitch processing in the brainstem. *Journal of Neurolinguistics*, *23*(1), 81–95.

Lee, G., Shin, D.-J., & Garcia, M. T. M. (2019). Perception of lexical stress and sentence focus by korean-speaking and spanish-speaking l2 learners of english. *Language Sciences*, *72*, 36–49.

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.

Li, J. J., & Grigos, M. (2018). The effect of l1 prosody in the perception and production of non-native lexical stress. In (pp. 898–902). <https://doi.org/10.21437/SpeechProsody.2018-181>

Li, X., To, C. K. S., & Ng, M. L. (2017). Effects of l1 tone on perception of l2 tone-a study of mandarin tone learning by native cantonese children. *Bilingualism: Language and Cognition*, *20*(3), 549–560.

Liburd, T. (2014). *Investigating prediction in l2 morpho-syntax: A visual world study* (PhD thesis). University of Pittsburgh.

Linck, J. A., Osthus, P., Koeth, J. T., & Bunting, M. F. (2014). Working memory and second language comprehension and production: A meta-analysis. *Psychon. Bull. Rev.*, *21*(4), 861–883.

Linzen, T., & Jaeger, T. F. (2016). Uncertainty and expectation in sentence processing: Evidence from subcategorization distributions. *Cognitive Science*, *40*(6), 1382–1411.

Lozano-Argüelles, C. (2020). *Predicting the end: Monolilnguals, l2 learners and interpreters’ use of prosody to predict word endings*.

Lozano-Argüelles, C., Sagarra, N., & Casillas, J. V. (2019). Slowly but surely: Interpreting facilitates l2 morphological anticipation based on suprasegmental and segmental information. *Bilingualism: Language and Cognition*, 1–11.

Malins, J. G., & Joanisse, M. F. (2010). The roles of tonal and segmental information in mandarin spoken word recognition: An eyetracking study. *Journal of Memory and Language*, *62*(4), 407–420.

Marull, C. (2017). Second language processing efficiency: Experience and cognitive effects on l2 morphosyntactic integration and anticipation. In *BUCLD 41: Proceedings of the 41st boston university conference on language development* (pp. 466–480). Cascadilla Press Somerville, MA.

Mirman, D. (2016). *Growth curve analysis and visualization using r*. CRC press.

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.

Nakamura, C., Arai, M., & Mazuka, R. (2012). Immediate use of prosody and context in predicting a syntactic structure. *Cognition*, *125*(2), 317–323.

Nuria Sagarra, N. R., Cristina Sanz. (2016). Eye-tracking reveals study abroad benefits on l2 morphosyntactic processing. Second Language Research Forum. Columbia University.

Ortega-Llebaria, M. (2006). Phonetic cues to stress and accent in spanish. In *Selected proceedings of the 2nd conference on laboratory approaches to spanish phonetics and phonology* (pp. 104–118). Cascadilla Proceedings Project Somerville, MA.

Ortega-Llebaria, M., Gu, H., & Fan, J. (2013). English speakers’ perception of spanish lexical stress: Context-driven l2 stress perception. *Journal of Phonetics*, *41*(3-4), 186–197.

Ortega-Llebaria, M., & Prieto, P. (2007). Disentangling stress from accent in spanish: Production patterns of the stress contrast in deaccented syllables. *Amsterdam Studies in the Theory and History of Linguistic Science Series 4*, *282*, 155.

Ortega-Llebaria, M., & Prieto, P. (2009). Perception of word stress in castilian spanish. *Phonetics and Phonology: Interactions and Interrelations*, *306*, 35.

Otten, M., & Van Berkum, J. J. (2009). Does working memory capacity affect the ability to predict upcoming words in discourse? *Brain Research*, *1291*, 92–101.

Pakulak, E., & Neville, H. J. (2010). Proficiency differences in syntactic processing of monolingual native speakers indexed by event-related potentials. *Journal of Cognitive Neuroscience*, *22*(12), 2728–2744.

Perdomo, M., & Kaan, E. (2019). Prosodic cues in second-language speech processing: A visual world eye-tracking study. *Second Language Research*, 0267658319879196.

Pozzan, L., Gleitman, L. R., & Trueswell, J. C. (2016). Semantic ambiguity and syntactic bootstrapping: The case of conjoined-subject intransitive sentences. *Language Learning and Development*, *12*(1), 14–41.

R Core Team. (2019). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>

Rehrig, G. L. (2017). *Acoustic correlates of syntax in sentence production and comprehension* (PhD thesis). Rutgers University-School of Graduate Studies.

Reichle, R. V., Tremblay, A., & Coughlin, C. (2016). Working memory capacity in l2 processing. *Probus*, *28*(1), 29–55.

Roll, M., & Horne, M. (2015). A neurolinguistic study of south swedish word accents: Electrical brain potentials in nouns and verbs, *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. (2011). Activating without inhibiting: Left-edge boundary tones and syntactic processing. *Journal of Cognitive Neuroscience*, *23*(5), 1170–1179.

Roll, M., Söderström, P., & Horne, M. (2013). Word-stem tones cue suffixes in the brain. *Brain Research*, *1520*, 116–120.

Sagarra, N. (2017). Longitudinal effects of working memory on l2 grammar and reading abilities. *Second Language Research*, *33*(3), 341–363.

Sagarra, N., & Casillas, J. V. (2018). Suprasegmental information cues morphological anticipation during l1/l2 lexical access. *Journal of Second Language Studies*, *1*(1), 31–59.

Sagarra, N., & Herschensohn, J. (2010). The role of proficiency and working memory in gender and number agreement processing in l1 and l2 spanish. *Lingua*, *120*(8), 2022–2039.

Salverda, A. P., Kleinschmidt, D., & Tanenhaus, M. K. (2014). Immediate effects of anticipatory coarticulation in spoken-word recognition. *Journal of Memory and Language*, *71*(1), 145–163.

Sanz, C., Lin, H.-J., Lado, B., Stafford, C. A., & Bowden, H. W. (2016). One size fits all? Learning conditions and working memory capacity in ab initio language development. *Applied Linguistics*, *37*(5), 669–692.

Saslow, M. (1967). Effects of components of displacement-step stimuli upon latency for saccadic eye movement. *Josa*, *57*(8), 1024–1029.

Schremm, A., Söderström, P., Horne, M., & Roll, M. (2016). Implicit acquisition of tone-suffix connections in l2 learners of swedish. *The Mental Lexicon*, *11*(1), 55–75.

Schwab, S., & Dellwo, V. (2016). The use of the odd-one-out task in the study of the perception of lexical stress in spanish by german-speaking listeners?

Schwab, S., Giroud, N., Meyer, M., & Dellwo, V. (2020). Working memory and not acoustic sensitivity is related to stress processing ability in a foreign language: An erp study. *Journal of Neurolinguistics*, *55*, 100897.

Serafini, E. J., & Sanz, C. (2016). Evidence for the decreasing impact of cognitive ability on second language development as proficiency increases. *Studies in Second Language Acquisition*, *38*(4), 607–646.

Sereno, J. A., & Lee, H. (2015). The contribution of segmental and tonal information in mandarin spoken word processing. *Language and Speech*, *58*(2), 131–151.

Soto-Faraco, S., Sebastián-Gallés, N., & Cutler, A. (2001). Segmental and suprasegmental cues for lexical access in spanish. *Journal of Memory and Language*, *45*, 412–432.

Söderström, P., Horne, M., & Roll, M. (2015). Using tonal cues to predict inflections. *Working Papers 55. 2015 Linguistics Lund University*, 91.

Söderström, P., Roll, M., & Horne, M. (2012). Processing morphologically conditioned word accents. *The Mental Lexicon*, *7*(1), 77–89.

Tanner, D., & Van Hell, J. G. (2014). ERPs reveal individual differences in morphosyntactic processing. *Neuropsychologia*, *56*, 289–301.

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods*, *37*(3), 498–505.

Van Leussen, J.-W., & Escudero, P. (2015). Learning to perceive and recognize a second language: The L2LP model revised. *Frontiers in Psychology*, *6*, 6–12. <https://doi.org/10.3389/fpsyg.2015.01000>

Vickie, Y. Y., & Andruski, J. E. (2010). A cross-language study of perception of lexical stress in english. *Journal of Psycholinguistic Research*, *39*(4), 323–344.

Weber, A., Grice, M., & Crocker, M. W. (2006). The role of prosody in the interpretation of structural ambiguities: A study of anticipatory eye movements. *Cognition*, *99*(2), B63–B72.

Wicha, N. Y., Bates, E. A., Moreno, E. M., & Kutas, M. (2003). Potato not pope: Human brain potentials to gender expectation and agreement in spanish spoken sentences. *Neuroscience Letters*, *346*(3), 165–168.

Wiener, S., & Turnbull, R. (2016). Constraints of tones, vowels and consonants on lexical selection in mandarin chinese. *Language and Speech*, *59*(1), 59–82.

Ye, Z., & Zhou, X. (2008). Involvement of cognitive control in sentence comprehension: Evidence from erps. *Brain Research*, *1203*, 103–115.

Zhu, X., & Wang, C. (2015). The oxford handbook of chinese linguistics. In W. S. Wang & C. Sun (Eds.). Oxford University Press.