



Open Library of Humanities

Perceptual sensitivity to stress in native English speakers learning Spanish as a second language

Ramsés Ortín, University of Texas Rio Grande Valley, Texas, United States of America, ramses.ortin@utrgv.edu

Miquel Simonet, University of Arizona, Arizona, United States of America, simonet@arizona.edu

Second language (L2) learners of Spanish whose first language (L1) is English tend to find Spanish lexical stress patterns difficult to acquire. This study investigates whether such difficulty derives, at least in part, from an obstacle encountered during perceptual processing: reduced perceptual sensitivity to stress distinctions. Participants were adult L1 English L2 Spanish learners of various proficiency levels. The experiment was a categorical matching task with triads of auditory stimuli minimally contrasting in stress (target) or segmental composition (baseline), an ABX task. The results show that learners were more accurate in the baseline condition than in the target condition, suggesting reduced perceptual sensitivity to stress relative to other contrasts. The reduction in accuracy, however, was restricted to trials in which matching items were not adjacent, further suggesting an obstacle with phonological processing in working memory rather than perceptual categorization. The default stress processing routines of L1 English L2 Spanish learners, optimized for their L1 (not their L2), may be responsible for their acquisitional obstacles with this feature of the Spanish language.



1. Introduction

Spanish phonology distinguishes words such as *bebe* [ˈbeβe] ‘she drinks’ and *bebé* [beˈβe] ‘baby,’ *libro* [ˈliβro] ‘book’ and *libró* [liˈβro] ‘freed, rescued.’ Words in such minimal pairs are distinguished solely on the basis of lexical stress, manifested via suprasegmental acoustic information, such as intensity, segment duration, and pitch (Ortega-Llebaria & Prieto, 2011). In Spanish, the first syllable in *bebe* [ˈbeβe] tends to be longer and higher in intensity than the first syllable in *bebé* [beˈβe], and the second syllable tends to be shorter and lower in intensity in *bebe* [ˈbeβe] than in *bebé* [beˈβe]. Syllables that receive such concentration of acoustic energy are referred to as being stressed relative to the surrounding syllables, which are unstressed. Minimal pairs such as *trusty-trustee* and *insight-incite* demonstrate that, like Spanish, English also possesses contrastive lexical stress (Cutler, 2005; Giegerich, 1992, pp. 179–207; Hualde, 2005, pp. 220–251). The present study investigates the processing of stress of English speakers learning Spanish as a second language with a focus on assessing their perceptual sensitivity to stress contrasts.

It has been reported that second language (L2) learners of Spanish whose first language (L1) is English seem to find Spanish stress relatively difficult to acquire (Beaudrie, 2007; Face, 2005; Kim, 2020; Ortega-Llebaria, Gu, & Fan, 2013; Ortín & Simonet, 2022; Romanelli & Menegotto, 2015; Romanelli, Menegotto, & Smyth, 2015; Saalfeld, 2012). This is typically noticed when such learners are taught the orthographic conventions of Spanish, which mark phonological stress with an acute accent mark on some words, as in *libró* [liˈβro] ‘freed.’ Since English does not mark stress in its spelling, it is not surprising that this marked feature of Spanish orthography attracts the attention of teachers and learners (Beaudrie, 2007, 2017). Whether a Spanish word is spelled with an accent mark or not depends on a set of simple rules. Importantly, the successful application of the spelling rules concerning Spanish stress depends on knowing *a priori* the location of stress in the word (Hualde, 2005, pp. 246–252). Beaudrie (2007) speculates that it is not the spelling rules *per se* that make Spanish orthographic stress difficult to acquire. What L1 English learners seem to struggle with is identifying the stressed syllable in the first place, that is, distinguishing it from the other syllables in the word. Beaudrie (2007) hypothesizes that the nature of the acquisitional obstacle L1 English L2 learners of Spanish face in the case of Spanish stress marking is perceptual, not orthographic.

Since the two languages of (emergent) bilinguals seem to share a representational space and the features of the two languages find themselves in constant interaction, phonetic and phonological obstacles in L2 acquisition are thought to be largely dependent on the sound structure of the learners’ L1 (Best & Tyler, 2007; Colantoni, Steele, & Escudero, 2015; Escudero, 2005; Flege, 1995; Flege, Aoyama, & Bohn, 2021; Flege & Bohn, 2021; Simonet, 2016; van Leussen & Escudero, 2015). For instance, English has a contrast between voiced and voiceless obstruents in word-final position, such as in *bad* and *bat*, that German lacks (Jessen & Ringen, 2002). In English, vowels preceding a voiced obstruent (as in *bad*) are longer than those preceding

a voiceless obstruent (as in *bat*) (Chen, 1970). L1 German L2 learners of English are known to merge the voicing-induced vowel duration difference in their English productions, effectively transferring a feature of German into their English (Smith, Hayes-Harb, Bruss, & Harker, 2009). L1 phonological transfer into the L2 is ubiquitous. This observation would initially suggest that the difficulties L1 English learners of L2 Spanish have with Spanish lexical stress may be due to some type of cross-linguistic influence from their L1. Lexical stress, however, is contrastive in both Spanish and English. Since stress is available to English speakers in their L1, what could be the *locus* of the obstacle L1 English speakers encounter when learning L2 Spanish stress? Or, in other words, why would learners experience a perceptual obstacle with a feature of their L2 they should be able to transfer from their L1?

1.1 Review of the literature

1.1.1 Perceptual processing and phonological representation

The present study explores one aspect of the difficulty L1 English learners of L2 Spanish seem to face when learning Spanish stress patterns: relatively weak perceptual processing routines concerning stress, likely inherited from their L1. The weak stress processing routines that affect L1 English L2 Spanish learners would, we hypothesize, manifest themselves as a diminished perceptual sensitivity to stress contrasts. In this study, we present evidence suggesting that this population indeed demonstrates a reduced perceptual sensitivity to stress.¹

The perception of speech involves a number of processes, from encoding the acoustic signal into discrete prelexical units (e.g., phonetic categories, phonemes, syllables) to ultimately accessing the mental lexicon and selecting a given lexical entry (Cutler, 2012; Dahan & Magnuson, 2006; Diehl, Lotto, & Holt, 2004; Samuel, 2011, 2020). Speech perception is influenced by the phonological and lexical representations in the listener's mind—that is, perception is modulated by the prelexical and lexical units available to the listener. Speech perception is not a purely bottom-up process (Cutler, 2012; Samuel, 2011, 2020). At the initial stages of perceptual processing, listeners are likely to depend on acoustic information whereas later processing stages are more likely to be influenced by the units in mental representation, including phonemic

¹ Empirical findings suggest that English speakers' perceptual sensitivity to stress is 'reduced' relative to that of speakers of some other languages with contrastive stress, such as Spanish, but not to that of speakers of languages that lack contrastive stress, such as French (Dupoux et al., 2001; Ortín & Simonet, 2022; Peperkamp et al., 2010). It is in this context that we use words such as 'reduced,' 'diminished,' 'deadened,' 'muted,' and 'weak.' Obviously, one could express this difference across language groups differently. For instance, one could say that Spanish speakers have 'heightened' perceptual sensitivity to stress relative to English speakers, thus placing English speakers in the unmarked position and Spanish speakers in the marked position. This would indeed avoid expressing the state of English speakers as a deficit, but it is not obvious that this would change any relevant aspect of the main observation. The observation is that the phonological structure of a speaker's main language leads them to demonstrate a perceptual sensitivity to stress that, on a spectrum, varies from 'deadened' to 'heightened' or from 'less' to 'more.' The location, on the spectrum, we consider to be unmarked is ultimately arbitrary and theoretically uninteresting.

contrasts, and less so by acoustic detail (Mann, 1986; Pisoni, 1973; Pisoni & Tash, 1974; Werker & Logan, 1985). Pisoni (1973), for instance, distinguished between two processing stages, or memory codes, in perceptual discrimination: auditory and phonetic. According to Pisoni (1973), the auditory code, available at the very early stages of processing, retains detailed acoustic information. At this stage, listeners can discriminate sounds on the basis of even minor acoustic detail. The rich auditory code, however, fades away rapidly and only some information is retained for further processing in subsequent stages. The phonetic memory code is available once sounds have been categorized, and it lasts longer than the auditory code; at this stage, discrimination is made on the basis of mental representations (abstract phonetic categories).

Werker and Logan (1985) distinguished between three processing stages: acoustic, phonetic, and phonemic. During acoustic processing, listeners discriminate stimuli “on the basis of any acoustic variability between individual exemplars” (Werker & Logan, 1985, p. 43). Listeners, however, are less likely to discriminate stimuli based on acoustic variability when the stimuli are separated by a longer interstimulus interval (Pisoni, 1973) or in procedures that have high working-memory demands, such as the ABX (or AXB) (Best, McRoberts, & Goodell, 2001; Carney, Widin, & Viemeister, 1977) or the sequence-recall tasks (Dupoux, Peperkamp, & Sebastián-Gallés, 2001). In the phonemic processing stage, listeners are likely to accurately discriminate stimuli only if they are phonemically contrastive in their L1 and they can classify them “according to the phonological categories used to contrast meaning in their native language” (Werker & Logan, 1985, p. 43). At this stage, abstraction has taken place and listeners no longer retain, in working memory, detailed acoustic information; discrimination at this stage is, therefore, driven by abstract mental representation. Whether there are two or three processing stages is not important for our purposes. The important point is that perceptual discrimination is a multifaceted process and that at least some of its aspects are modulated by available representational units more than by the detailed acoustic information in the input.

The present paper utilizes the ABX paradigm to investigate the perceptual processing of stress of L1 English L2 Spanish learners because this paradigm allows us to assess the relative dependence of listeners on acoustic information *versus* phonological competence (or mental representation), and thus it taps into at least two of the stages identified in Werker and Logan (1985) and Pisoni (1973). In the ABX paradigm, participants are presented with three auditory stimuli in a sequence, and they are asked to match the last stimulus in the triad to either the first or the second. The triads are designed so that there is only one possible match, either A (the first) or B (the second). In this paradigm, the matching stimuli may be adjacent (ABB, BAA) or not (ABA, BAB). Since they are separated by a relatively long period of time and an intervening stimulus, listeners may not compare nonadjacent stimuli on the basis of acoustic similarity alone but must rely on a representational unit available to them in phonological working memory. Matching adjacent stimuli, on the other hand, is more likely to be influenced by acoustic detail.

As such, this experimental paradigm allows us to identify what perceptual units are (and are not) available in phonological working memory in different populations: These would be the ones that may be used to accurately match nonadjacent stimuli (Best et al., 2001; Dupoux, Pallier, Sebastián-Gallés, & Mehler, 1997).

1.1.2 Perceptual sensitivity to stress contrasts across languages

In some languages, such as Spanish and English, stress is phonemically contrastive; in others, such as French and Finnish, it is not. The functional role of stress in the phonological structure of a language has been found to modulate the degree of perceptual sensitivity to stress demonstrated by its speakers (Dupoux, Pallier, Sebastián-Gallés, & Mehler, 1997, 2001; Lin, Wang, Idsardi, & Xu, 2014; Peperkamp, Vendelin, & Dupoux, 2010; Qin, Chien, & Tremblay, 2017). For instance, using an ABX paradigm, Dupoux et al. (1997) found that L1 French speakers showed a lower perceptual sensitivity to stress distinctions than L1 Spanish speakers. Arguably, the fact that lexical stress is not a feature of French phonology ‘encourages’ French listeners to ignore suprasegmental differences at the word level during perceptual processing, even in nonwords (i.e., sensitivity to stress is deadened); or, alternatively, L1 Spanish speakers are ‘trained’ (by their L1) to become sensitive to stress (i.e., sensitivity to stress is heightened). In addition to their ABX experiment, Dupoux et al. (1997) ran an AX (same-different) task. For this experiment, listeners were presented with two adjacent auditory stimuli, and they were asked to indicate whether the stimuli were different (e.g., [‘bedapi]-[be’dapi]) or whether they were identical (e.g., [be’dapi]-[be’dapi]). Concerning the AX task, Dupoux et al. (1997) found that the response accuracy for the French listeners was almost at ceiling (i.e., they displayed high accuracy). Together, the ABX and the AX results of Dupoux et al. (1997) suggest that French listeners demonstrate relatively low perceptual sensitivity to stress, but only in experimental paradigms that tap into phonological working memory (likely, the phonemic processing stage of Werker & Logan, 1985), not in those for which listeners may rely on auditory (or acoustic) processing (Pisoni, 1973; Werker & Logan, 1985). In other words, French listeners may have “trouble representing and storing in working memory accent patterns that are otherwise accurately perceived (Dupoux et al., 1997, p. 416).”

A more recent body of literature revisited this phenomenon with a new experimental paradigm: The sequence-recall task (Dupoux et al., 2001; Peperkamp et al., 2010). In the sequence-recall task, listeners are taught to associate a pair of auditorily presented nonwords to two different keys on a computer keyboard; for instance, on a numeric keypad, [‘numi] may be associated with ‘1’ and [nu’mi] with ‘2.’ Listeners are later presented with trials with a varied number of auditory stimuli (four, five, or six nonwords in a sequence), and they are asked to type the order of presentation of the items. Thus, for the sequence [nu’mi]-[nu’mi]-[‘numi]-[nu’mi], a listener would be expected to type 2212 in this order. The sequence-recall task is very cognitively

demanding, much more so than the ABX task. As such, this task taps fully into phonological working memory rather than acoustic perception.

The sequence-recall literature on the perception of stress has found that speakers of languages in which stress is contrastive (and thus unpredictable), such as Spanish, are very sensitive to stress differences, whereas speakers of languages that lack contrastive lexical stress, such as French, Finnish, Hungarian, and Korean, are not (Dupoux et al., 2001; Lin et al., 2014; Peperkamp et al., 2010). An interesting case is presented by Polish. In Polish, stress is not contrastive, but neither is it entirely predictable from the surface form of words: There is a small number of lexical exceptions to the stress-assignment rule. As a default, primary stress in Polish falls on the penultimate syllable of the word, and most suffixes do not alter this rule; that is, as suffixes are added to a stem, stress moves ‘to the right’ to stay on the penultimate syllable. Exceptionally, however, stress may fall on the antepenultimate syllable in Polish. Antepenultimate stress occurs mostly in borrowings and in learned words (i.e., words of Latin and Ancient Greek origin), and two inflectional suffixes used in past-tense verbal forms do not trigger stress displacement, *-śmy* (1st plural) and *-ście* (2nd plural); thus, such forms display antepenultimate stress. Exceptions to the Polish stress rule are either morphologically conditioned or due to borrowing (Domahs et al., 2012). In a sequence-recall experiment, it was found that Polish speakers are less sensitive to stress distinctions than Spanish speakers but more so than French speakers (Peperkamp et al., 2010). These findings suggest that the phonological role of stress in a language modulates the perceptual sensitivity to stress differences in its speakers. Presumably, the contrastive use of stress in a language, such as Spanish, induces its speakers to become perceptually sensitive to stress. This, in turn, suggests that lexical stress is part of the phonological competence of Spanish speakers, but not that of French speakers. The findings regarding Polish suggest that, beyond contrastivity (or the lack of it), the lexical statistics of a given phonological feature can lead to situations in which speakers’ perceptual sensitivity to the feature is ‘intermediate.’ Perceptual sensitivity to stress lies, therefore, on a spectrum.

Some of the research with the sequence-recall paradigm has focused on bilingual and L2 learner populations (Dupoux et al., 2010, 2008; Lin et al., 2014; Ortín & Simonet, 2022; Qin et al., 2017). For the most part, these studies have found that L1 speakers of a language that lacks contrastive lexical stress show (relatively) deadened sensitivity to stress even when they speak a L2 that has contrastive stress. This suggests that a low sensitivity to stress, induced by the structure of the phonology of the L1, could present an acquisitional obstacle in cases where a L2 makes contrastive use of lexical stress and the L1 does not. What about, however, L1 English L2 Spanish learners? Since stress is lexically contrastive in English, why would these learners seem to have any difficulties with Spanish stress?

1.1.3 Stress in Spanish and English: Phonological patterns and perceptual processing

In Spanish, stress is primarily cued by suprasegmental acoustic information: duration, intensity, and pitch (Ortega-Llebaria & Prieto, 2011). Duration seems to be the most robust acoustic correlate of stress in this language (Torreira, Simonet, & Hualde, 2014). Pitch, on the other hand, seems to be a correlate of phrasal accent—intonation, a phonological property of phrases, not words. When a word is in nuclear phrasal position and it receives a pitch accent, its stressed syllable will likely be of longer duration, higher intensity, and higher pitch than those surrounding it—unless the pitch accent is a low tone, in which case pitch remains low during the stressed syllable. On the other hand, when the word is not in nuclear position and it lacks a pitch accent (or the pitch accent is instantiated by a low tone), duration may remain the only reliable acoustic correlate of stress (Ortega-Llebaria & Prieto, 2011). If a word is in pre-nuclear position and it receives a rising-tone pitch accent, its tonal peak is likely to be realized on the post-stressed syllable, and thus higher pitch becomes a property of the post-stressed syllable rather than the stressed one while longer duration remains on the stressed syllable.

Unlike English, Spanish does not have phonological vowel reduction—the same inventory of vowel categories is available in both stressed and unstressed positions (Nadeu, 2014). In English, stress is primarily cued by vowel quality—compare, for instance, the pronunciation of the first two syllables in *photograph* ['fouərəɡræf] and *photographer* [fə'thɑɡræfə]. This is not to say that suprasegmental acoustic cues are not used to signal stress in English—in fact, the same acoustic correlates that are used in Spanish are also available in English (Morrill, 2012; Ortega-Llebaria et al., 2013; Ortega-Llebaria & Prieto, 2011; Sluijter & van Heuven, 1996a, 1996b). As a matter of fact, some English minimal pairs are not distinguished by vowel reduction, but solely by suprasegmental cues such as duration and intensity—e.g., *insight-incite*, *trusty-trustee*, *permit* (noun)-*permit* (verb), *forbear* (noun)-*forbear* (verb). These cases are evidence of the existence of the type of contrastive stress that is also found in Spanish. Nevertheless, it is fair to say that, in English, stress distinctions that do not rely on vowel quality alternations are rare.

Although low in both languages, the functional load of stress (i.e., the number of minimal pairs dependent on stress) is higher in Spanish than it is in English (Cutler & Pasveer, 2006). In English, contrastive stress occurs mostly between nouns and verbs—e.g., *permit* (noun) is a paroxytone while *permit* (verb) is an oxytone. English has few stress-based minimal pairs. In Spanish, on the other hand, stress minimal pairs occur frequently within the same part of speech (e.g., *canto* 'I sing' ['kaŋto], *cantó* 's/he sang' [kaŋ'to]). In fact, Spanish stress-based minimal pairs abound within verbal paradigms (Hualde, 2005, pp. 228–233).

Cutler and colleagues (Cutler, Norris, & Sebastián-Gallés, 2004; Cutler & Pasveer, 2006) identified an additional difference between Spanish and English stress patterns that concerns the phenomenon of embedded words. Words may include other words embedded within their

phonological form—for instance, *sea* is embedded in *secret*, and *may* is embedded in *maple*. During lexical access, embedded words are spuriously activated—*sea* may be briefly activated in a process that ultimately selects *secret* (Cutler, 2012). It turns out that Spanish has many more embedded words than English does but, crucially, only if the computation ignores stress. For instance, when considering only segments, *casa* ['kasa] ‘house’ is embedded in *casado* [ka'saðo] ‘married.’ If the computation requires a match in stress configuration, however, there is no overlap: The first syllable in *casa* is stressed, whereas it is unstressed in *casado*. It turns out that, in Spanish, the number of embedded words is reduced drastically when the computation is sensitive to stress relative to when it is not (Cutler et al., 2004; Cutler & Pasveer, 2006). In English, on the other hand, the difference between including or excluding stress in the computation turns out to be negligible. This suggests that, for speakers of Spanish, it might be crucial to represent stress in the mental representation of the phonological form of words, but perhaps not for English speakers.

Some experimental evidence indeed suggests that Spanish speakers exploit stress to resolve lexical competition during spoken word recognition while English speakers do not (Cooper, Cutler, & Wales, 2002; Cutler, 1986; Soto-Faraco, Sebastián-Gallés, & Cutler, 2001; van Donselaar, Koster, & Cutler, 2005). With a cross-modal fragment priming paradigm, Soto-Faraco et al. (2001) found that, in L1 Spanish speakers, the Spanish sequence ['prinθi] primed *príncipe* ['prinθipe] ‘prince’ but inhibited *principio* [prin'θipjo] ‘beginning’—a mismatch in stress configuration between the prime and the target led to lexical inhibition. In English, on the other hand, mismatches of this sort do not cause lexical inhibition (Cooper et al., 2002; Cutler, 1986), and Cutler (1986) found that members of stress-based minimal pairs such as *trustee* and *trusty* primed each other’s semantic associates.

The findings and observations discussed in this section suggest that differences in functional load, patterns of perceptual (lexical) processing, and (perhaps) phonological representation pertaining to stress between Spanish and English could lead to differences in overall perceptual sensitivity to stress between speakers of these languages. Consequently, this could cause acquisitional obstacles to L1 English L2 Spanish learners. The present study investigates perceptual sensitivity to stress in L1 English L2 Spanish learners. Our study explores this with an experimental paradigm that taps into a processing stage that relies on phonological working memory and not auditory perception.

1.1.4 Perceptual processing of stress in L1 English L2 Spanish learners

L1 English speakers learning Spanish as a L2 appear to find the acquisition of Spanish stress somewhat difficult (Beaudrie, 2007; Face, 2005; Kim, 2020; Ortega-Llebaria et al., 2013; Ortín & Simonet, 2022; Romanelli & Menegotto, 2015; Romanelli et al., 2015; Saalfeld, 2012). Researchers have explored the behavior of L1 English L2 Spanish listeners in terms of their

perceptual identification of stressed syllables in multisyllabic nonwords (Beaudrie, 2007; Kim, 2020; Ortega-Llebaria et al., 2013; Romanelli & Menegotto, 2015; Romanelli et al., 2015). For instance, in two studies, learners heard auditory stimuli, nonwords, such as [se'mapa] or [sema'pa] and were asked to identify the nonword they had heard from among a list of visually-presented options, such as <semapa> and <semapá> (Romanelli & Menegotto, 2015; Romanelli et al., 2015). In another study, learners heard nonword auditory stimuli, such as [ta'xira], and were asked to identify the stressed syllable in the nonword by selecting from among a list of corresponding visual materials in which the stressed syllable was indicated in capital letters: <TAgira>, <taGIra>, <tagiRA> (Beaudrie, 2007). L1 Spanish controls, as predicted, had no problem selecting the visual item that corresponded to the auditory stimulus (or identifying the stressed syllable), but L2 learners were found to perform rather poorly in these experiments.

Another relevant identification experiment is that of Ortega-Llebaria et al. (2013). For this study, L1 English L2 Spanish listeners and L1 Spanish controls were played target auditory stimuli in which suprasegmental acoustic cues to stress had been orthogonally manipulated and target words appeared on a variety of phrasal conditions, such as declarative sentences and reporting clauses. Learners and controls were found to rely on acoustic cues differently, and acoustic-cue reliance was further dependent on phrasal context, which conditions the presence of pitch accents and their shape. The findings of Ortega-Llebaria et al. (2013) suggest that, at the early stages of perceptual processing, L1 English L2 Spanish learners may depend on slightly different acoustic correlates of stress than L1 Spanish speakers, at least in some circumstances. Ortega-Llebaria et al.'s (2013) listening experiment was an identification experiment in which listeners were given a closed set of response options and heard a single auditory stimulus per trial. The minimal pair used in Ortega-Llebaria et al. (2013), [ma'ma]-['mama], consisted of two pronunciation variants of the same lexical item, meaning 'mother.' To participate in this task, listeners were "told that the word *mama* could be pronounced either as [ma'ma] or ['mama], and were instructed to press the keyboard space bar when they heard the oxytone word [ma'ma] (Ortega-Llebaria et al., 2013, p. 191)."

The point has been made that these studies hardly demonstrate that L1 English L2 Spanish learners have trouble with the perceptual processing of Spanish stress *per se* (Ortín & Simonet, 2022). These studies indeed suggest that such learners may have trouble identifying stressed syllables explicitly or mapping nonword auditory stimuli to visual renderings of such nonwords. What these studies suggest, therefore, is that L2 learners may have limited *phonological awareness* of stress (Gillon, 2017), but not necessarily compromised perceptual processing of stress. Phonological awareness relies on explicit phonological knowledge (a form of metalinguistic knowledge), not implicit competence. A difficulty concerning perceptual processing would indeed be demonstrated if listeners were shown to have poor discrimination of functioning stress contrasts; for instance, if a listener were shown to not 'hear' the difference between *caso* ['kaso]

‘case’ and *casó* [ka'so] ‘married’ (that is, to not discriminate these two words or to process them as homophones), one could certainly speak of reduced perceptual processing of stress. For listeners to accurately respond to experimental trials in explicit tasks such as those in Romanelli and Menegotto (2015), Romanelli et al. (2015), Beaudrie (2007), and Ortega-Llebaria et al. (2013), they must rely on explicit awareness of Spanish phonology and Spanish spelling—that is, they must be able to understand terms such as ‘stress’ and ‘syllable’ as used in task instructions. One could conceive of participants—perhaps L1 Spanish speakers who happen to be illiterate—who have no problems discriminating between [ˈkaso] and [ka'so] but whose explicit awareness of stress might be rather limited. An illiterate Spanish speaker, for instance, would have linguistic, tacit knowledge of stress that would allow them to distinguish [ˈkaso] from [ka'so] but might not be able to find words to explain the precise linguistic feature that leads to the phonemic contrast or might not be able to participate in a task whose instructions include words such as ‘stress’ and ‘syllable.’

Four studies have explored the perceptual processing of stress of L2 English L1 Spanish learners using experimental paradigms that tap into implicit phonological competence rather than explicit phonological or metalinguistic awareness (Kim, 2020; Ortín & Simonet, 2022; Saalfeld, 2012; Sagarra & Casillas, 2018). Of these, two studies focused on perceptual identification in a meaningful context (Kim, 2020; Sagarra & Casillas, 2018), one focused on perceptual discrimination, also in a meaningful context (Saalfeld, 2012), and one was concerned with the phonological processing of stress contrasts in a sequence-recall task with nonwords (Ortín & Simonet, 2022). In Kim (2020), participants were played single-presentation auditory stimuli, such as *paso* [ˈpaso] ‘I pass (first person, present tense)’ and *pasó* [pa'so] ‘s/he passed (third person, past tense),’ and they were asked to indicate the agent of the verb: *yo* ‘I’ (first person) or *él* ‘he’ (third person). As a group, L1 English L2 Spanish learners were accurate only about 63% of the time (Kim, 2020). Sagarra and Casillas (2018) ran an eye-tracking study in which listeners heard auditory stimuli, such as *firma* [ˈfirma] ‘signature’ and *firmó* [fir'mo] ‘s/he signed,’ and were asked to match each stimulus with one of two visually-presented options, such as <firma> and <firmó>. The eye-tracking data showed that the L1 Spanish speakers and a group of advanced L2 learners anticipated the word ending from the suprasegmental acoustic information present in the first syllable. A group of intermediate L2 learners, on the other hand, did not show evidence of anticipation. Together, the findings in Kim (2020) and Sagarra and Casillas (2018) suggest that L1 English L2 learners of Spanish encounter an acquisitional obstacle when confronted with Spanish stress, at least at the initial and intermediate stages of learning. These studies suggest that the acquisitional obstacle these learners encounter concerns implicit phonological competence and affects auditory lexical access, but they do not demonstrate that the obstacle has to do with reduced perceptual sensitivity to stress *per se*—i.e., compromised phonological processing routines specifically concerned with stress.

Saalfeld (2012) conducted a perceptual discrimination study with full Spanish sentences. In this experiment, listeners were auditorily presented with three sentences in a sequence, where the first and last were different from each other and the second matched either the first or the third. Participants were asked to indicate the matching sentences. Sentences were relatively long—such as *hable* [a'βle] *con la profesora después de clase* ‘speak with the teacher after class’ and *hablé* [a'βle] *con la profesora después de clase* ‘I spoke with the teacher after class,’ and they differed only on the stress configuration of the main verb. The response patterns of L2 learners showed them to be at chance in this experiment. Since it used long carrier phrases, Saalfeld’s (2012) task was arguably very cognitively demanding; also, it utilized real Spanish words and sentences. Unlike identification tasks, discrimination tasks do not rely fully on lexical competency; participants do not need to speak the language of the experiment, and experiments may use nonwords. When participating in Saalfeld’s task, however, listeners likely activated multiple lexical items and grammatical structures, and had doubts knowing where to focus their attention.

With a large sample of L1 English L2 Spanish learners, Ortín and Simonet (2022) conducted a sequence-recall task using bisyllabic nonwords. Learners, but not L1 Spanish controls, were found to be more accurate when recalling sequences whose stimuli differed in their segmental composition, a baseline condition (e.g., [tuki]-[tuki]- [tupi]-[tupi]-[tupi]-[tuki]), than sequences that differed in their stress configuration (e.g., [nu'mi]-[numi]-[numi]-[nu'mi]-[numi]-[nu'mi]). Like Saalfeld (2012), Ortín and Simonet (2022) assessed L2 learners’ perceptual sensitivity to stress contrasts with a discrimination paradigm. Unlike Saalfeld (2012), Ortín and Simonet (2022) used nonwords, which allowed for the assessment of the perceptual processing of stress without the need to activate lexical and grammatical competency. Arguably, both studies showed that L1 English L2 Spanish learners have a deadened perceptual sensitivity to stress. This was shown, in both cases, in tasks that demand a very high working-memory load (Ortín & Simonet, 2022; Saalfeld, 2012). The present study revisits this issue with a discrimination paradigm that imposes lower working-memory demands than the experiments in both Saalfeld (2012) and Ortín and Simonet (2021), but higher than those in Ortega-Llebaria et al. (2013). This begins to chart the limits of the phenomenon.

The present experiment differs from previous ones in that it analyzes implicit perceptual sensitivity to stress contrasts in L1 English L2 Spanish learners independently from their lexical and grammatical knowledge—that is, in tasks that do not rely on the processing of words or on explicit phonological awareness or metalinguistic knowledge (Beaudrie, 2007; Kim, 2020; Ortega-Llebaria et al., 2013; Romanelli & Menegotto, 2015; Romanelli et al., 2015; Sagarra & Casillas, 2018). The present experiment also differs from previous studies in that it explores implicit perceptual sensitivity at a processing stage that is, arguably, intermediate between the low-level auditory perception mode tapped into in Ortega-Llebaria et al. (2013) and the high-level

processing mode tapped into in high-demand working-memory tasks focused on discrimination (Ortín & Simonet, 2022; Saalfeld, 2012).

1.2 The present study

This study reports on an ABX perceptual discrimination experiment with nonword auditory stimuli based on the design used in Dupoux et al. (1997). As in Dupoux et al. (1997), the presentation trials in our study differed as a function of whether contrasts relied on consonantal manipulations ([ne¹ðapi]-[be¹ðapi]-[ne¹ðapi]) or, rather, manipulations to the stress configuration of the nonwords ([¹beðapi]-[be¹ðapi]-[¹beðapi]). The consonant condition was our baseline condition: a within-subject, within-task control for phonological working-memory capacity. Our focus was the stress condition. We collected perceptual discrimination data from a large group of L1 English L2 Spanish learners and a small comparison group of L1 Spanish speakers who happened to be L2 learners of English. The crucial question we addressed in our study was whether listeners' perceptual sensitivity, in a within-subjects design, was lower in the target condition than in the baseline condition. We did not compare the two groups of listeners directly because such a comparison is not informative. By conducting within-subjects analysis, we focused on the role of stress relative to a baseline.

The main goal of the present study was to investigate whether L1 English L2 Spanish learners display reduced, deadened perceptual sensitivity to stress (relative to a baseline) in a task that taps into phonological working-memory, the ABX—a task that can reveal the strength of listeners' mental representation of stress. The strength of listeners' mental representations of stress was directly assessed by comparing their response patterns in trials in which the matching stimuli were adjacent (ABB, BAA) with those in trials in which they were not (ABA, BAB). Participants could not match the stimuli on the basis of low-level acoustic detail in any case, since the stimuli in a trial were recorded from different talkers, but the working-memory demands were still arguably different for adjacent and nonadjacent matching stimuli (Best, McRoberts, & Sithole, 1988). This allowed us to chart the limits of L1 English L2 Spanish learners' compromised perceptual sensitivity to stress, shown elsewhere (Ortín & Simonet, 2022; Saalfeld, 2012). Also, since we recruited L2 learners from a variety of proficiency levels, we were able to investigate the extent to which perceptual sensitivity to stress distinctions may change with increased experience in the L2, if at all (Ortín & Simonet, 2022).

2. Method

2.1. Participants

A total of 86 young adults were sampled from two populations to participate in a perceptual discrimination study. We had a small ($N = 10$) group of L1 Spanish speakers, who acted as controls, and a large ($N = 76$) group of L1 English speakers who, at the time of the study, were

learning Spanish as a L2 in a school setting, in a tertiary-education program. Recruitment and study participation took place on the campus of the University of Arizona, in Tucson, Arizona.

The control group was composed of graduate students who had been raised in a Spanish-speaking country as monolingual Spanish speakers (four of them in Mexico, six in Spain), learned English as a L2 in school starting at approximately age 12, and moved to the US as adults, approximately at age 22, to pursue a graduate degree. The L2 group was composed of undergraduate students who had been raised in the southwestern US (Arizona or Southern California) as English-speaking monolinguals. At the time of the study, they were enrolled in Spanish language classes in college, and their use of Spanish was restricted to that setting. The participants in the L2 group were recruited from a range of Spanish language classes, including first-, second-, and third-year courses.

The participants, including the controls, completed three Spanish language proficiency tests and a linguistic profile questionnaire in addition to the perceptual discrimination experiment that forms the basis of the present study. The chosen questionnaire was the Bilingual Language Dominance Profile, or BLP (Gertken, Amengual, & Birdsong, 2014). The proficiency tests comprised two cloze tests, of which one was a passage and the other a collection of independent sentences (Martínez García, 2016). A vocabulary-size test, the LexTale-Esp, was also administered (Izura, Cuetos, & Brysbaert, 2014).

The BLP is a self-report questionnaire that provides data regarding each of the two languages of a bilingual in four modules: personal language history, personal language use, self-assessed linguistic proficiency, and language attitudes. The answers to each of these modules are numeric and yield one score per language. These scores are then used to calculate a *dominance* (or linguistic orientation) score by subtracting one from the other. To the extent that such score deviates from zero, it indicates a dominance of one of the two languages over the other. In our study, negative values indicate preferred orientation towards (or dominance in) Spanish, and positive values are indicative of dominance in English. We were very careful not to include any early-childhood bilinguals in our sample, and the BLP was useful in that regard. **Table 1** reports the descriptive statistics pertaining to the participants' dominance.

Table 1 also reports the descriptive statistics pertaining to the participants' Spanish proficiency scores. The LexTale-Esp is a test used to assess Spanish language proficiency through an assessment of vocabulary size (Izura et al., 2014). The version of the test we used includes 60 Spanish words and 30 nonwords, and the participants' task was to decide on the lexicality of the word forms (word, not a word). The resulting score is calculated by subtracting 2 for each false alarm and adding 1 for each correct hit. The maximum possible score is 60, and negative scores are possible. A higher value is suggestive of a larger vocabulary. The two cloze tests, adapted from Martínez García (2016), were administered via an online form. One was a passage from which 20 words had been replaced for blanks, and the other had 30 sentences from which one word per sentence

had been replaced for a blank. In both tests, participants were given a closed set of options, four, by means of a drop-down menu. The resulting proficiency scores were obtained by adding the correct responses, for a maximum of 20 points for the passage cloze test and 30 points for the sentence cloze test. Higher values are suggestive of higher grammatical proficiency in Spanish.

	N	Dominance		Sentence		Passage		Vocabulary	
		M	SD	M	SD	M	SD	M	SD
Native Speakers	10	-129.7	(18)	29	(0.8)	17.9	(0.7)	57.4	(1.7)
L2 Learners	76	145.4	(24)	10	(3.7)	6.2	(2.3)	0.3	(6.5)
1 st year	27	165	(18)	7.81	(2.3)	5.81	(1.7)	-2.04	(6.6)
2 nd year	24	144	(20)	9.79	(3.6)	6.21	(2.4)	0.38	(6.1)
3 rd year	25	125	(16)	12.84	(3.3)	6.64	(2.7)	2.76	(6.1)

Table 1: Descriptive statistics of three Spanish proficiency tests and a language profile questionnaire. The dominance score was calculated from the answers to a language profile questionnaire, the sentence and passage scores resulted from two cloze tests, and the vocabulary scores were obtained with a vocabulary-size test. L2 learners are further grouped as a function of their year of enrollment in college Spanish (1st year, 2nd year, 3rd year).

For the L2 learners only, the proficiency and dominance scores were normalized (*z*-scored) so that they could be compared across measurement scales. In this subset of the data (which excluded the controls), most (but not all) of the normalized scores were associated with each other: passage-sentence, $r = .30$ (95% CI [.08, .49], $p = .009$), passage-vocabulary, $r = .14$ (95% CI [-.09, .35], $p = .234$), sentence-vocabulary, $r = .44$ (95% CI [.24, .60], $p < .0001$), vocabulary-dominance, $r = -.30$ (95% CI [-.49, -.08], $p = .009$), passage-dominance, $r = -.14$ (95% CI [-.36, .08], $p = .215$), and sentence-dominance, $r = -.52$ (95% CI [-.67, -.34], $p < .0001$). The scores were not found to reliably measure the same construct: Cronbach's $\alpha = .64$. Subsequently, a multivariate analysis of variance (MANOVA) explored the potential effects of year of enrollment (first, second, third) on the normalized scores. Year of enrollment predicted about half the variance in the variables, Wilk's $\lambda = .48$, $F(8,140) = 7.85$, $p < .0001$. The test also found that, at the .05 significance threshold for p , three of the metrics were statistically modulated by year of enrollment: dominance, $F(2,73) = 17.17$, $p < .0001$, sentence cloze test, $F(2,73) = 12.07$, $p < .0001$, passage cloze test, $F(2,73) = 0.84$, $p > .05$ [.44], and vocabulary size, $F(2,73) = 3.79$, $p < .05$ [.027].

2.2. Instrument

Our instrument is a conceptual replication (an adaptation) of one of the experiments in Dupoux et al. (1997), an ABX perceptual discrimination task that made use of 12 CVCVCV nonword quadruplets differing in three possible contrast conditions: consonant, stress, and redundant. While this task measures perceptual discrimination, it is appropriate to think of it as a matching

task: Listeners are played three items and asked to match the two that are identical on some phonological property. In the consonant condition, stimuli differed on a consonant contrast shared by both Spanish and English, such as /b/-/n/ (e.g., ['beðapi]-['neðapi]). This was our control condition, and all participants were predicted to discriminate the items in such trials relatively easily. This condition provided us with a working-memory baseline. Contrasting consonants were found as onsets to any of the three syllables in the word forms. In the stress condition, stimuli in a trial differed in their stress configuration, and stress could fall on any of the first two syllables in the word form (e.g., ['beðapi]-[be'ðapi]). This was our target condition, and the L2 learners were hypothesized to find it more difficult to discriminate than the consonant condition. We included a third condition as our exclusion criterion: The redundant condition. In this condition, stimuli in a trial differed with respect to both consonantal and stress configuration (e.g., ['beðapi]-[ne'ðapi]). This condition is expected to be particularly easy, or at least as easy as the consonant condition. We included it as our exclusion criterion to filter out participants who might not have been engaging with the task or might have been distracted during their participation. Before we explored the rest of the data, we calculated the proportion of trials in the redundant condition responded to accurately by each participant. We set out to exclude the data of any participant who scored below 90% correct responses in this condition. The 86 participants whose data were retained for the present study satisfied this criterion. The redundant condition was not explored any further; it was dropped from all analyses.

The stimuli were presented in the form of three auditory items per trial: A triad. The first and second items (A, B) were always different word forms, which contrasted as a function of one of three key contrast conditions, as explained above. The third item in the triad (X) was always categorically identical to either the first (A) or the second (B) item. Participants were asked to indicate, by pressing a button, whether the last item matched the first one (ABA, BAB) or, rather, the second one (ABB, BAA). Trials in which the target matched the first item were assigned to the *primacy* condition, and those in which the target matched the second item were assigned to the *recency* condition. Comparisons in the primacy condition were expected to be more challenging than those in the recency condition, particularly in cases in which phonological processing is compromised, as matching stimuli are relatively distant from each other and an additional stimulus is heard between the two matching items (Best et al., 1988).

In total—that is, including the redundant condition—the task had 288 trials. Trials were counterbalanced so that each of the items in the quadruplets had the opportunity to appear as the first item three times, one for every possible contrast found in the second item: stress, contrast, and redundant. This resulted in 12 possible combinations for each quadruplet (4 word forms \times 3 contrasts). Since the third item, the target, could match either the first or second one in the triad, all the possible binary combinations of each quadruplet appeared twice, one for every possible correct order: primacy condition [ABA, BAB], and recency condition [ABB, BAA]. The design, therefore, was as follows: 4 quadruplets \times 3 contrasts \times 12 combinations \times 2 orders = 288

trials. The order of presentation of trials was randomized for each participant, and participants were given the option to take a break every 24 trials. The stimulus onset asynchrony (SOA) was set at 500 ms. Recall that, for the analysis, we retained only 192 of the 288 trials (or 2/3), since we excluded all the trials that instantiated the redundant contrast condition.

The auditory stimuli were recorded from three L1 Spanish-speaking talkers, two men and one woman. The three talkers, born and raised in Spain, grew up speaking only Spanish and acquired English as young adults. The talkers were recorded in the US, where they resided at the time of the study. Whereas they used English daily, they remained dominant in Spanish, and they used the latter daily as well. Recordings were made with a Marantz PMD660 digital recorder and a Shure SM10A head-mounted, dynamic microphone. The recordings were digitized at a sampling rate of 44.1 kHz and 16-bit quantization, and then normalized for intensity level. In the trials, talker order was kept constant. The stimuli from the two male talkers were assigned to the first and second slots (A, B), and the stimuli from the female talker was consistently played as the third item in the triads (X). This ensured that even the matching word forms were acoustically different, as they had been produced by different talkers (of different genders). Thus, auditory comparisons performed by the participants needed to be based on categorical comparisons rather than on acoustic memory; They required the ability to abstract away from phonetic detail, thus activating phonological representations to the extent that they are available to the participants. Rather than assessing perceptual discrimination, it is perhaps accurate to say that the ABX paradigm measures phonological processing in working memory. Relatively speaking, this is particularly true of the primacy condition, where matching items are not adjacent.

Participants used Sony MDR-7502 headphones, and they sat inside a sound-attenuated booth in front of a computer running the experiment from a Python script in PsychoPy (Peirce, 2007; Peirce, Gray, Simpson, MacAskill, Höchenberger, Sogo, Kastman, & Lindeløv, 2019). Participants were told that they would listen to “three Spanish false words” (that is, ‘pseudowords produced by native speakers of Spanish’) and were instructed to decide, as fast and as accurately as possible, whether the last item in the triad matched the first or the second one. They were told that there were no catch trials; in all trials, the third item matched either the first or the second one, and there was always a correct response. Instructions to all participants were given in English. Participants were asked to respond to each triad by pressing one of two activated keys on a Cedrus RB response pad (Cedrus Corporation, San Pedro, California). Responses were coded as a function of their accuracy and latency (ms). Latency was measured from the acoustic onset of the third auditory item, the target. The experiment was preceded by a practice session with 10 trials with items under identical conditions as the experimental trials but with different word forms.

Table 2 reports some of the descriptive statistics pertaining to some of the acoustic information of the auditory stimuli used in the study. The table shows the mean, minimum, and maximum of each of the acoustic cues for each of the three talkers. The acoustic observations include the duration of the first and second vowels in each of the nonce words (all of them trisyllabic), the

mean intensity across the length of each of the target vowels, and the mean f_0 (in Hz) across the length of each of the target vowels. A cursory examination of **Table 2** suggests that all three talkers behave similarly in terms of the acoustic correlates used in their productions, with clear differences in both duration and mean f_0 between stressed and unstressed syllables and perhaps a negligible effect of intensity. A spreadsheet in the supplementary materials lists all the triads used in the study, and it shows all the nonce words chosen. A cursory examination of the spreadsheet shows that a variety of vowels and consonants were used, which suggests the findings are generalizable across a range of sounds and potential acoustic correlates of stress, including segmental effects.

Value	Talker	Stress	Duration (ms)		Intensity (dB)		f_0 (Hz)	
			V1	V2	V1	V2	V1	V2
Mean	X	[ˈcvcvcv]	118	82	70	62	215	170
		[cvˈcvcv]	93	128	65	68	188	203
	B	[ˈcvcvcv]	123	91	70	64	135	102
		[cvˈcvcv]	94	127	65	70	104	120
	A	[ˈcvcvcv]	104	74	70	66	138	115
		[cvˈcvcv]	73	136	66	69	112	126
Minimum	X	[ˈcvcvcv]	95	51	68	49	196	154
		[cvˈcvcv]	70	106	57	59	109	182
	B	[ˈcvcvcv]	100	66	66	55	118	90
		[cvˈcvcv]	50	97	60	65	98	110
	A	[ˈcvcvcv]	78	43	66	59	118	105
		[cvˈcvcv]	45	102	58	66	102	118
Maximum	X	[ˈcvcvcv]	139	97	72	70	238	184
		[cvˈcvcv]	130	162	71	71	210	221
	B	[ˈcvcvcv]	152	110	72	71	144	111
		[cvˈcvcv]	121	157	71	72	110	128
	A	[ˈcvcvcv]	119	95	72	71	153	126
		[cvˈcvcv]	92	191	71	71	121	141

Table 2: Descriptive statistics (mean, minimum, maximum) of some acoustic correlates of stress as a function of talker (A [male], B [male], X [female]) and stress configuration for auditory stimuli used in the stress condition (stressed on first syllable [ˈcvcvcv], stressed on second syllable [cvˈcvcv]) only. The acoustic correlates include duration of the first vowel (stressed in [ˈcvcvcv], and unstressed in the [cvˈcvcv]), duration of the second vowel (unstressed in [ˈcvcvcv], and stressed in [cvˈcvcv]), mean intensity across the length of the first vowel, mean intensity across the length of the second vowel, mean f_0 across the length of the first vowel, and mean f_0 across the length of the second vowel.

2.3. Data analysis

Since groups were far from being balanced in size, the data from the control and experimental groups were analyzed separately. An L1 Spanish control group was included merely to ensure the adequacy of the instrument, as L1 Spanish speakers had been studied before with this paradigm, albeit not with these particular auditory items (Dupoux et al., 1997). Regarding the L1 Spanish controls, we expected to replicate the findings in Dupoux et al. (1997) with our design and materials, which would suggest that our task was reliable and amenable for use with other populations. Our study, therefore, focused on the L2 learners, and this explains why our L2 group was much larger than our control group.

Data were explored in two main steps. The first statistical exploration consisted of analyses of accuracy rates and response times as a function of two factorial conditions: type of *contrast* (consonant, stress) and *order* of stimuli (primacy [ABA, BAB], recency [ABB, BAA]). Repeated-measures analyses of variance were used, one per dependent variable and participant sample. The significance, or α , threshold was set at $p = .05$. For each participant, accuracy rates (proportion of correct responses) and mean response times (ms) were obtained, for each experimental condition, by averaging across relevant trials—48 trials per condition and participant. Aggregating the data resulted in four accuracy values and four timing values per participant for a total of 344 accuracy observations (4 conditions \times 86 participants) and 344 timing observations (4 conditions \times 86 participants).

Participants' responses were excluded from the averaging process according to the following criteria. First, all trials responded to faster than 250 ms and slower than 2250 ms were filtered out. Trials responded to faster than 250 ms are likely to capture responses to stimuli not in the trial (perhaps the preceding trial), and they are likely to be false alarms or spurious responses. Trials responded to slower than 2250 ms are likely to capture the consequence of a distracted decision process. Our criteria resulted in the exclusion of only statistical outliers, and not large numbers of observations. In the analysis of the timing data, we further excluded all incorrect responses, as they would be uninformative. These exclusion criteria resulted in the exclusion of 40 observations from the control data set (1920–40 = 1880, 2% excluded) and of 353 observations from the learner data set (1459–353 = 14239, 2.5% excluded). The number of observations excluded from the timing analysis is obviously larger, and this is reported in **Table 3**.

The second statistical exploration of the data set focused exclusively on the learners' perceptual behavior and investigated whether linguistic proficiency and dominance—as measured by the data we gathered from each participant—were associated with perceptual discrimination patterns. In other words, this exploration investigated whether an increase in proficiency and/or dominance—as measured by normalized dominance and proficiency scores—explained a significant amount of variance in the learners' phonological processing patterns. Linear regression was the basis for such explorations.

The proportion of correct responses (P) per condition and participant were subjected to a *logit* transformation, $\log(P/(1-P))$, prior to submission to inferential statistics, and response times in ms were subjected to a log transformation, $\log(\text{ms})$. This increased the normality of the distribution by reducing positive skewness, typical of response-time data. Data preparation was done with a collection of *R* (version 4.0.1) scripts (r-project.org), with package *tidyverse* (version 1.3.1) (tidyverse.org) (Wickham et al., 2019). Statistical data analyses were conducted in *Jamovi* (version 1.2.22) (The Jamovi Project, 2020), a free and open-source GUI for *R* (jamovi.org). Three *R* packages were used in the analyses: *afex* (version 0.27-2) (Singman, Bolker, Westfall, Aust, & Ben-Shachar, 2020), *emmeans* (version 1.4.7) (Lenth, 2018), and *esci* (version 0.9.1) (see Calin-Jageman & Cumming, 2019; Cumming, 2013; Cumming & Calin-Jageman, 2016).²

3. Results

Table 3 reports the descriptive statistics of the untransformed accuracy (P correct) and response timing (ms) data, respectively, as a function of Contrast (consonant, stress) and Order (primacy, recency). Data are further broken down by participant group. All inferential statistical explorations were conducted solely with the transformed dependent variables.

Metric	Group	Consonant						Stress					
		Primacy			Recency			Primacy			Recency		
		<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>
Accuracy	L2	3574	.94	.24	3591	.96	.21	3484	.72	.45	3590	.94	.24
	NS	471	.97	.16	475	.97	.17	461	.98	.15	473	.98	.14
Timing	L2	3362	984	294	3432	935	274	2493	1080	312	3360	975	269
	NS	459	957	286	461	935	310	450	1029	332	464	945	294

Table 3: Descriptive statistics of response accuracy (proportion of correct responses) and response timing (ms) in a perceptual matching experiment as a function of two within-subjects factors, Contrast (consonant, stress) and Order (primacy, recency), further divided as a function of two participant groups: native Spanish controls (NS), $N = 10$, and L2 learners of Spanish (L2), $N = 76$.

3.1 L1 Spanish speakers

The first analysis is concerned with the response accuracy data. The logit-transformed accuracy values were submitted to a $(2) \times (2)$ repeated-measures ANOVA with Contrast (consonant, stress) and Order (primacy, recency) as factors. None of the main effects were statistically significant at the predetermined threshold, that is, neither Contrast, $F(1,9) = .34, p > .05$ [.57], $\eta_p^2 = .037$,

² Instruments, synthetic data, and analysis scripts may be made available to readers interested in reproducing our analyses by e-mailing the corresponding author.

nor Order, $F(1,9) = .14, p > .05$ [.71], $\eta_p^2 = .015$, were significant. The interaction also failed to reach statistical significance, $F(1,9) = 1.88, p > .05$ [.203], $\eta_p^2 = .173$. The estimated marginal means and their 95% confidence intervals (CIs) are reported in **Table 4**. At this juncture, there is no positive evidence for us to reject the null hypothesis, and we conclude that neither type of contrast (baseline versus target condition), nor order of presentation significantly affect response accuracy in L1 Spanish speakers. Across conditions, L1 controls were fundamentally at ceiling in this task.

Contrast	Order	Accuracy (logit)		Timing (log)	
		<i>M</i>	95% CI	<i>M</i>	95% CI
Consonant	Primacy	5.28	[3.95, 6.60]	6.82	[6.74, 6.91]
	Recency	4.93	[3.61, 6.25]	6.79	[6.70, 6.88]
Stress	Primacy	4.45	[3.13, 5.77]	6.89	[6.81, 6.98]
	Recency	5.07	[3.75, 6.39]	6.81	[6.72, 6.90]

Table 4: Estimated marginal means and 95% CIs of response *accuracy* (logit) and response *timing* (log) as a function of Contrast (consonant, stress) and Order (primacy, recency). The data are from 10 native Spanish speakers serving as a control group.

The second analysis concerns the response times. The log-transformed latency values were also analyzed with a $(2) \times (2)$ repeated-measures ANOVA with both Contrast (consonant, stress) and Order (primacy, recency). At the predetermined significance threshold, the ANOVA found main effects of both Contrast, $F(1,9) = 6.68, p < .05$ [.0295], $\eta_p^2 = .426$, and Order, $F(1,9) = 13.97, p < .05$ [.0046], $\eta_p^2 = .608$, but not a Contrast by Order interaction, $F(1,9) = 3.619, p > .05$ [.0895], $\eta_p^2 = .287$. The estimated marginal means and their 95% CIs are reported in **Table 4**. On average, participants reacted faster in the consonant condition than in the stress condition, $M_{\text{diff}} = 0.04$ (0.02 SE), and they were slower to respond in primacy trials than in recency trials, $M_{\text{diff}} = 0.06$ (0.02 SE).

To summarize, L1 speakers of Spanish were found to be highly accurate when matching auditory stimuli in both consonant and stress trials, regardless of order of presentation. In terms of matching accuracy, they were at ceiling in all conditions. Responses were found to be slower in primacy trials than in recency ones, and they were also slower in the stress than in the consonant contrast conditions. We surmise that the L1 controls behaved as expected—that is, they behaved similarly to the Spanish-speaking participants in Dupoux et al. (1997). Note that our sample of L1 participants is rather small, which leads to a relatively imprecise experiment with relatively low informational value—this can be seen in the width of the 95% CIs.

3.2 L1 English L2 Spanish learners

We begin the analysis of the L2 learners' behavior by exploring the response accuracy data. As with the L1 controls, the logit-transformed accuracy values were submitted to a $(2) \times (2)$ repeated-measures ANOVA with Contrast (consonant, stress) and Order (primacy, recency) as factors. In this case, however, both main effects and their interaction were found to be statistically significant: Contrast, $F(1,75) = 48.60, p < .0001, \eta^2_p = .39$, Order, $F(1,75) = 52.74, p < .0001, \eta^2_p = .41$, and Contrast by Order interaction, $F(1,75) = 55.62, p < .0001, \eta^2_p = .43$. Planned pairwise comparisons are reported in **Table 5**. Overall, participants were less accurate with stress contrasts than with consonant contrasts, $M_{\text{diff}} = 1.14$ (0.16 SE), and they were more accurate in the recency condition than in the primacy condition, $M_{\text{diff}} = 1.20$ (0.17 SE). The interaction is due to the fact that the participants were much less accurate in the stress-primacy (sub)condition than in the other three (sub)conditions. On the one hand, order of presentation modulated response accuracy in stress trials but not in consonant trials; on the other, consonant and stress trials were different from each other in the primacy condition but not in the recency condition. The estimated marginal means and their 95% CIs are plotted in **Figure 1**, and **Figure 2** plots estimated marginal means, 95% CIs, and individual scores for the raw values (that is, for *accuracy* measured as *P* correct scores rather than logit units), for visual comparison. The results of the inferential tests allow us to reject the null hypothesis that, in the L2 data, experimental conditions do not affect response accuracy. Unlike the controls, we conclude, the learners were less likely to be accurate in stress trials than in consonant trials, but, interestingly, only when auditory stimuli were presented in the primacy order, not in the recency order. In other words, the participants were particularly prone to error when the target and its match differed in stress condition, and they were not adjacent.

Comparison									
Condition 1			Condition 2		df	<i>t</i>	<i>p</i> _{tuckey}	<i>M</i> _{diff}	<i>SE</i> _{diff}
Consonant	Primacy	–	Consonant	Recency	141.4	–1.15	.6575	–0.24	0.21
Stress	Primacy	–	Stress	Recency	141.4	–10.31	<.0001	–2.16	0.21
Consonant	Primacy	–	Stress	Primacy	141.8	10.09	<.0001	2.10	0.21
Consonant	Recency	–	Stress	Recency	141.8	0.89	.8081	0.19	0.21

Table 5: Planned pairwise comparisons pertaining to repeated measures ANOVA with response accuracy (logit) as dependent variable and Contrast (consonant, stress) and Order (primacy, recency) as factors. The data are from 76 L2 learners of Spanish.

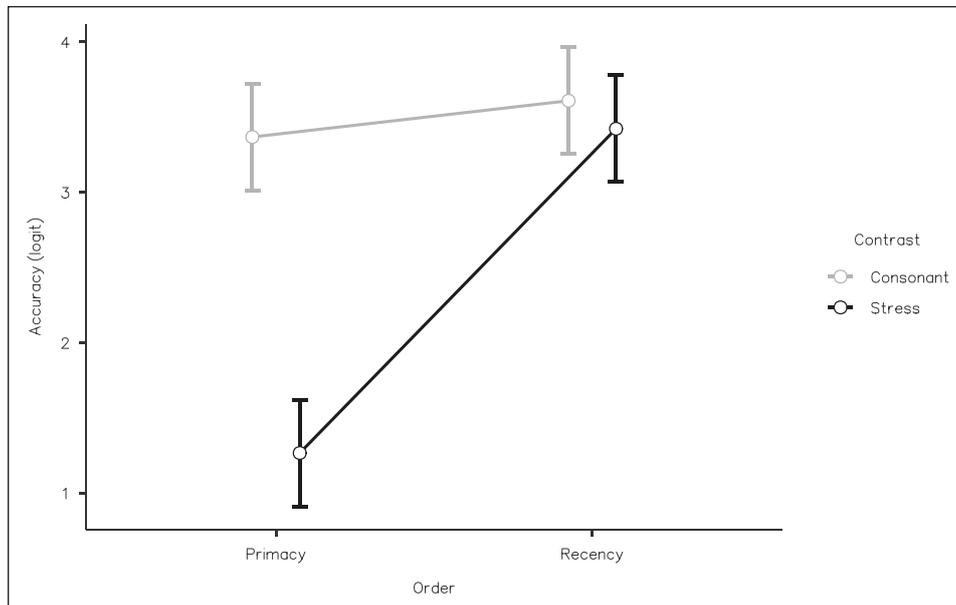


Figure 1: Estimated marginal means and 95% CIs of response accuracy (logit) as a function of Contrast (consonant, stress) and Order (primacy, recency). The data come from 76 L2 learners of Spanish.

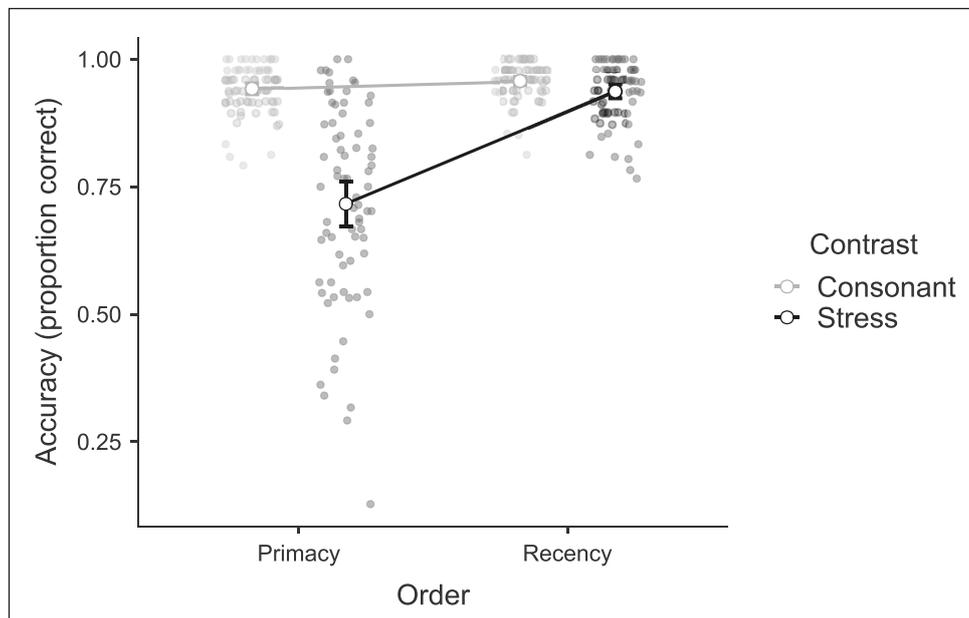


Figure 2: Estimated marginal means, 95% CIs, and individual observations of raw response accuracy (P correct) as a function of Contrast (consonant, stress) and Order (primacy, recency). The data come from 76 L2 learners of Spanish.

The second analysis is concerned with response timing. The log-transformed timing values were analyzed with a $(2) \times (2)$ repeated-measures ANOVA with both Contrast (consonant, stress) and Order (primacy, recency). The ANOVA found statistically significant effects of both main factors as well as a significant interaction between the two: Contrast, $F(1,75) = 150.34$, $p < .0001$, $\eta^2_p = .67$, Order, $F(1,75) = 105.30$, $p < .0001$, $\eta^2_p = .58$, and Contrast by Order interaction, $F(1,75) = 27.49$, $p < .0001$, $\eta^2_p = .27$. Planned pairwise comparisons are reported in **Table 6**. Estimated marginal means and their 95% CIs are plotted in **Figure 3**, and, for visual comparison, **Figure 4** plots estimated marginal means, 95% CIs, and individual observations for the raw values (that is, for timing in ms rather than log units). Participants were found to be slower to respond in stress trials than in consonant trials, $M_{\text{diff}} = 0.07$ (0.006 SE), and they were slower in the primacy than in the recency condition, $M_{\text{diff}} = 0.08$ (0.007 SE). The pattern one infers from the pairwise planned comparisons is that order of presentation affected both contrast types, but the effects of order were much larger in stress trials than in consonant trials. Additionally, participants were faster to respond in consonant trials than in stress ones, but this effect was much larger in the primacy than in the recency condition. In other words, the fastest condition was the consonant-recency (sub)condition, and the slowest one was the stress-primacy (sub)condition, with the other two (sub)conditions falling in between.

Comparison									
Condition 1			Condition 2		df	<i>t</i>	<i>p</i> _{tuckey}	<i>M</i> _{diff}	<i>SE</i> _{diff}
Consonant	Primacy	–	Consonant	Recency	131.0	5.61	<.0001	0.05	0.01
Stress	Primacy	–	Stress	Recency	131.0	11.44	<.0001	0.10	0.01
Consonant	Primacy	–	Stress	Primacy	144.4	–12.81	<.0001	–0.10	0.01
Consonant	Recency	–	Stress	Recency	144.4	–6.16	<.0001	–0.05	0.01

Table 6: Planned pairwise comparisons pertaining to repeated measures ANOVA with response timing (log) as dependent variable and Contrast (consonant, stress) and Order (primacy, recency) as factors. The data are from 76 L2 learners of Spanish.

To summarize, the L2 learners were found to be less accurate when responding to the stress contrast condition, but only in trials in which the matching stimuli were not adjacent. Additionally, as a group, the L2 learners were found to be slower to respond to stress trials than to consonant trials, particularly in the primacy condition—that is, in trials in which the matching stimuli were not adjacent.

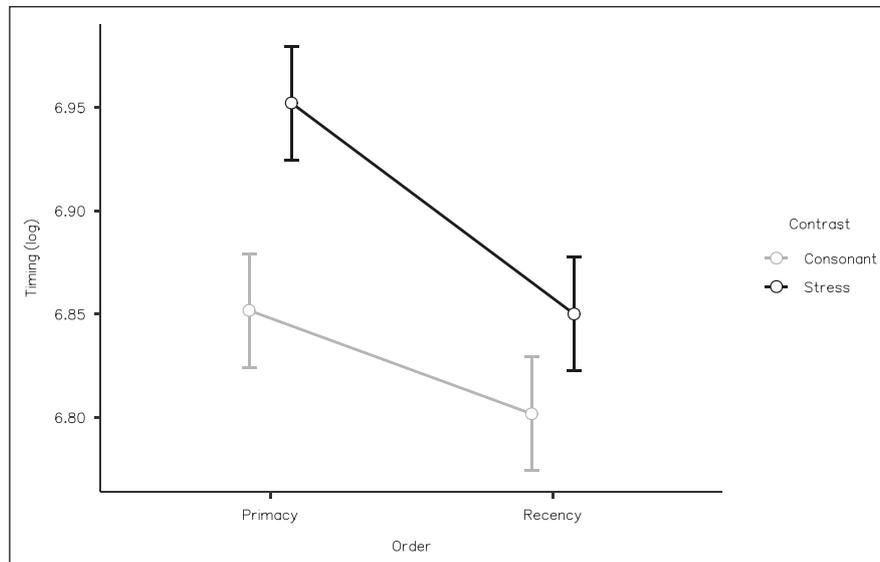


Figure 3: Estimated marginal means and 95% CIs of response timing (log) as a function of Contrast (consonant, stress) and Order (primacy, recency). The data come from 76 L2 learners of Spanish.

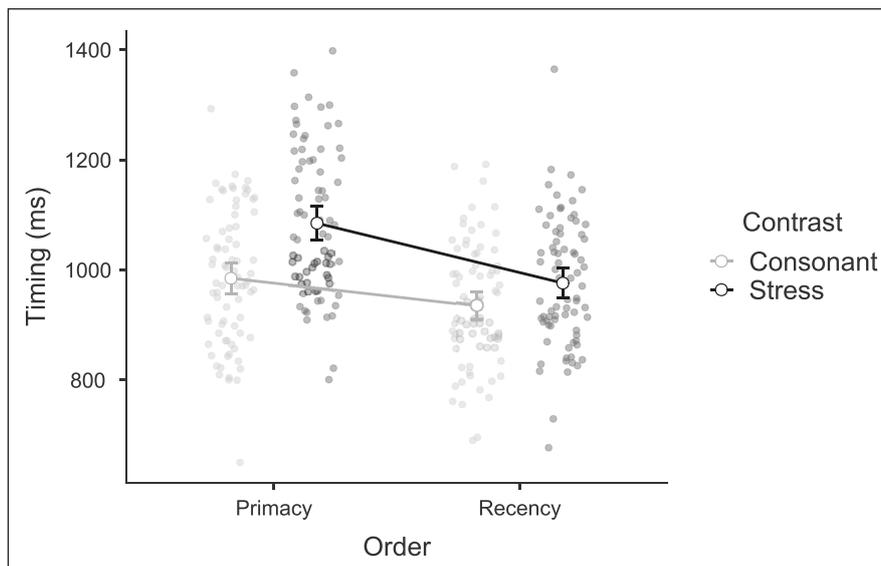


Figure 4: Estimated marginal means, 95% CIs, and individual scores of raw response timing (ms) as a function of Contrast (consonant, stress) and Order (primacy, recency). The data come from 76 L2 learners of Spanish.

3.3 The role of second language proficiency

Does perceptual sensitivity to stress heighten with increased L2 experience and/or proficiency? To address this question, we focused on the primacy condition, as the preceding analyses revealed that gains are likely only in this condition—relative to all other conditions, a drop in

perceptual matching accuracy of stress trials was concentrated in the primacy condition while other conditions revealed response accuracy to be at ceiling.

Note that the participants' accuracy scores in the stress-primacy condition were correlated with their scores in the consonant-primacy condition: $r = .37$, 95% CI [.16, .55], $p = .001$. In other words, participants' higher accuracy in the consonant condition was associated with higher accuracy in the stress condition. This likely reveals differences in working-memory capacity among individuals, and such differences are not particularly relevant here. For us to conclude that stress processing abilities—and not merely participation in ABX experiments with Spanish auditory stimuli—improves with increases in L2 proficiency in L2 learners, we must demonstrate that the difference between the target and the baseline condition decreases with increases in L2 proficiency or extended L2 experience. In other words, if we want to be able to say that something is affecting the processing of stress (in particular), we must be able to capture changes in the stress condition *relative* to a participant's own accuracy in the consonant condition, not their overall accuracy (in both conditions) in primacy trials. We therefore obtained a new metric.

To obtain a new dependent variable, we subtracted the participants' accuracy (logit) scores in the stress-primacy condition from those in the consonant-primacy condition to obtain a within-participant difference score, a contrast effect specific to the primacy condition. This became our dependent variable. We then analyzed whether such a contrast effect—the new dependent variable—was associated with any of the proficiency and experience predictors: the participants' dominance score, their sentence cloze test score, their passage cloze test score, and their vocabulary size test scores. The results of the correlation analyses are reported in **Table 7**. None of the proficiency and experience scores were found to be associated with the contrast effect in the primacy condition: The dependent variable. A linear regression model with accuracy difference as response and all four proficiency and experience indicators as predictors failed to reach statistical significance, $R^2 = .08$, $F(4,71) = 1.55$, $p > .05$ [.196]. We conclude that we were not able to find any robust evidence suggesting that perceptual discrimination of stress distinctions improves with increased Spanish proficiency or with increased experience in the language, at least not at the initial stages of learning, which are the stages that characterize our participant sample.

Predictor	Pearson's r	95% CI	P
Dominance Scores (BLP)	-.08	[-0.30, 0.15]	.49
Sentence Cloze Test Scores	-.19	[-0.39, 0.04]	.11
Passage Cloze Test Scores	-.00	[-0.23, 0.22]	.97
Vocabulary Size Test Scores	-.04	[-0.27, 0.18]	.70

Table 7: Correlation coefficients of contrast effect (i.e., by-participant accuracy logit scores in the consonant-primacy condition minus accuracy logit scores in stress-primacy condition) and four L2 experience and proficiency indicators, all of them normalized (z -scored). The data are from 76 L2 learners of Spanish.

4. Discussion

4.1 Summary of findings

We recruited a group of 10 L1 Spanish speakers, who acted as controls to verify the adequacy of the instrument, and a large group of 76 L2 learners of Spanish with English as their L1. The L1 Spanish speakers were at ceiling accuracy in all conditions; in fact, they were slightly more accurate in the stress condition than in the baseline condition. In terms of their response latencies, it was found that these participants were slightly faster when responding to trials differing in consonantal composition (baseline) than in trials differing in stress configuration, and they were faster in trials in which the matching stimuli were adjacent than in those in which they were not adjacent.

The L2 learners were found to be at ceiling when responding to three of the four conditions in the experiment. The only condition that negatively affected the learners' accuracy rates, relative to the others, was that in which matching stimuli were not adjacent and auditory stimuli differed in stress configuration, that is, the stress-primacy condition. In other words, learners' perceptual sensitivity to stress was relatively deadened only at the intersection of the stress and primacy conditions. In terms of their response latencies, the L2 learners were also particularly slow to respond to the stress-primacy condition relative to the other three conditions. Interestingly, there was no clear evidence that L2 proficiency or experience—as measured by two cloze tests, a vocabulary-size test, and a language profile questionnaire—modulated perceptual discrimination of stress distinctions in this participant sample in any direction.

It is important to conclude the analysis of the data with an estimation of the magnitude of the most relevant effects in the L2 sample. We focus here on the effects of contrast in the primacy condition, since L2 participants were almost at ceiling in the recency condition. In this order condition, learners were much less likely to be accurate in the stress (target) than in the consonant (baseline) trials, $M_{diff} = 2.102$, 95 % CI [1.71, 2.49], $r = .37$. When standardized, the effect was revealed to be very large, Cohen's $d_{avg} = 1.37$, 95% CI [1.09, 1.69]. This effect is plotted in **Figure 5**. In the same order condition (primacy), learners were slower to respond in the stress trials than in the consonant trials, $M_{diff} = 0.10$, 95 % CI [0.083, 0.118], $r = .802$, and, once again, this effect proved to be large, though not as much, Cohen's $d_{avg} = 0.82$, 95% CI [0.66, 1.00]. The effect is plotted in **Figure 6**. Note that the 95% CI in both figures are quite narrow, suggesting a precise and informative estimate of the population. The magnitude of the relevant effects in the L2 sample and, by extension, the population is certainly not negligible and must be accounted for.

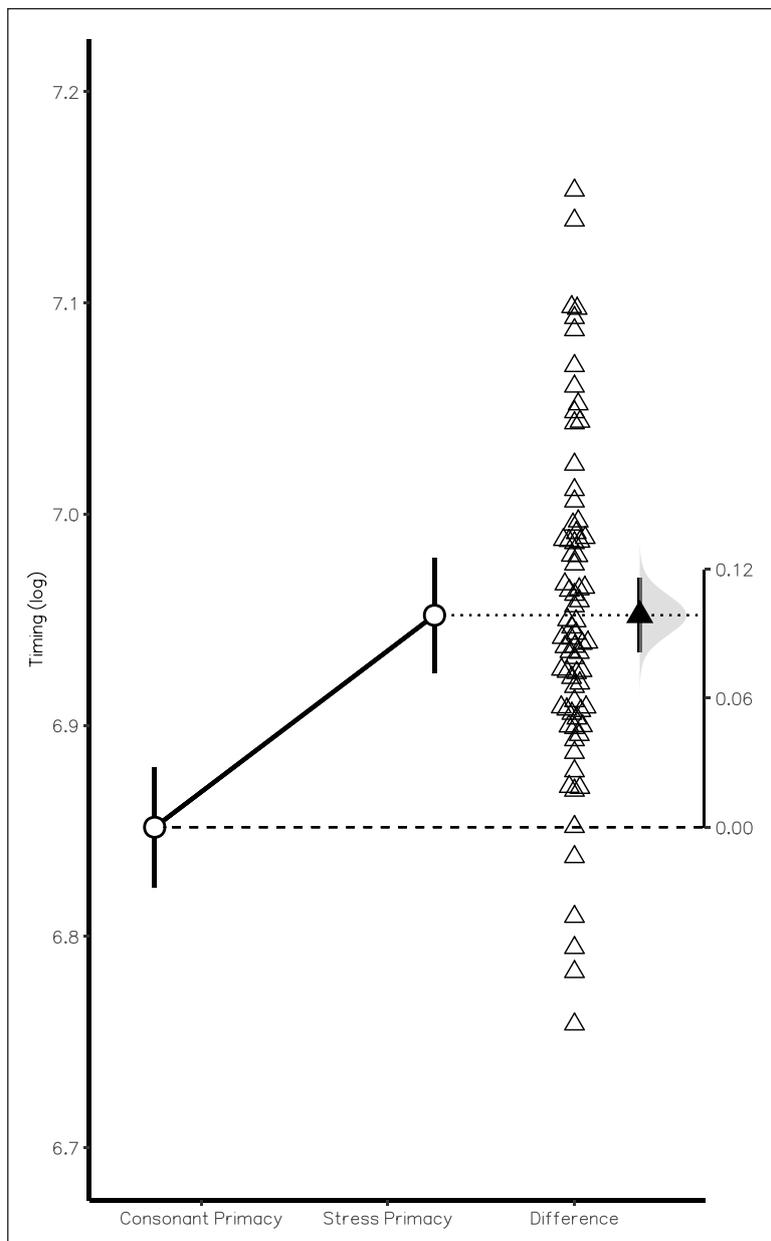


Figure 5: Paired effect of Contrast (consonant, stress) on response accuracy (logit) in the primacy condition in a sample of 76 L2 learners of Spanish. Error bars are 95% CI around each group mean. The left-most bar plots the group mean for the consonant-primacy condition, the center bar plots that of the stress-primacy condition, and the right-most bar plots the mean of the by-participant differences (and its 95% CI), that is, a distribution of the differences between the consonant-primacy and stress-primacy conditions for each subject. The white triangles plot difference values for each participant, that is, the individual difference scores (76 observations). The scale on the right-hand side shows the size of the paired effect in logit units. The two y-axis scales (left-hand, right-hand) are in the same units, but the one on the right hand represents only the difference between the two conditions, expressing one condition as zero and the other as distance from zero.

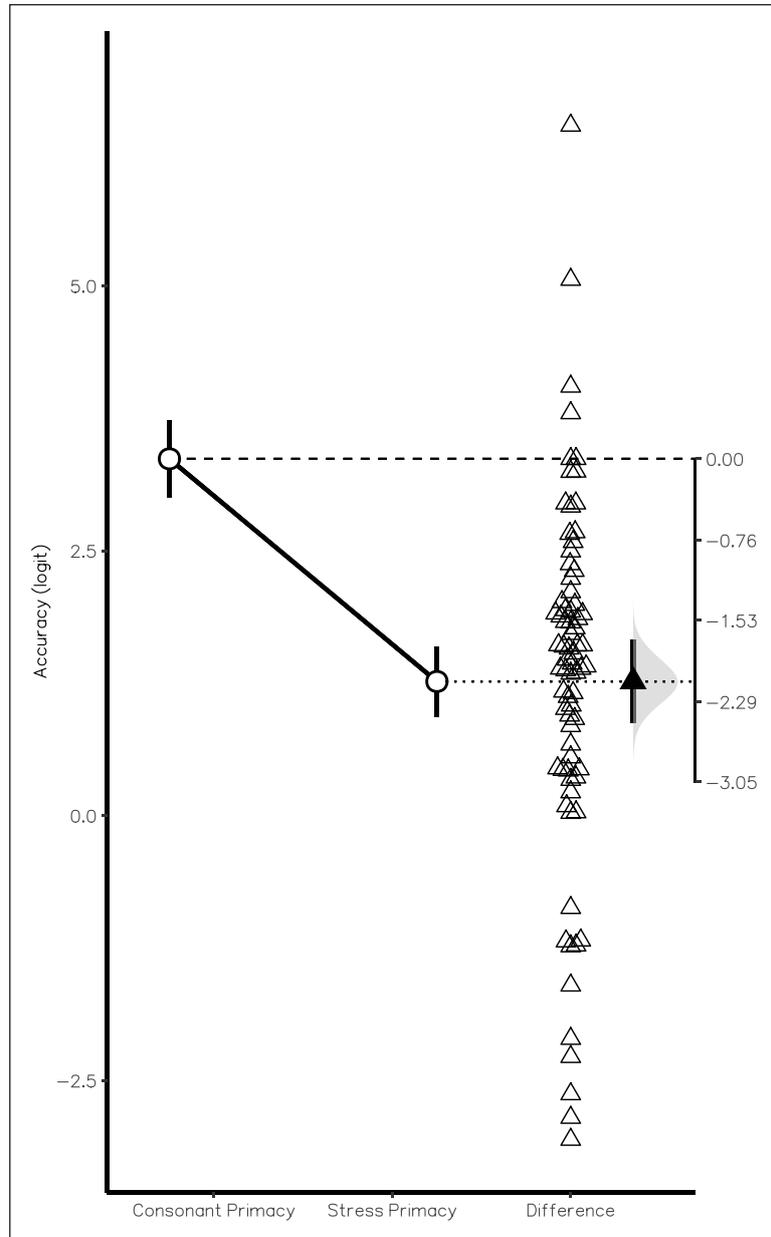


Figure 6: Paired effects of Contrast (consonant, stress) on response timing (log) in the primacy condition in a sample of 76 L2 learners of Spanish. Error bars are 95% CI around each group mean. The left-most bar plots the group mean for the consonant-primacy condition, the center bar plots that of the stress-primacy condition, and the right-most bar plots the mean of the by-participant differences (and its 95% CI), that is, a distribution of the differences between the consonant-primacy and stress-primacy conditions for each subject. The white triangles plot difference values for each participant, that is, the individual difference scores (76 observations). The scale on the right-hand side shows the size of the paired effect in log units. The two y-axis scales (left-hand, right-hand) are in the same units, but the one on the right hand represents only the difference between the two conditions, expressing one condition as zero and the other as distance from zero.

4.2 Interpretation and implications

The discussion of the results of the present study focuses exclusively on the findings pertaining to the L1 English L2 Spanish learners and, particularly, on the fact that these participants' perceptual sensitivity to stress was found to be deadened, muted in the primacy condition—that is, when the matching auditory stimuli were not adjacent. In terms of response accuracy, the L2 learners were at ceiling in the baseline condition (in both the recency and primacy trials) and in the stress-recency condition (where the matching items were adjacent), but they were not in the stress-primacy condition (where the matching items were not adjacent). The response time data are fundamentally in line with the patterns revealed by the accuracy data; the discussion thus centers around the accuracy data.

The interest in L1 English L2 Spanish learners' perceptual sensitivity to stress originates from the fact that Spanish graphically marks the presence of stress over the prominent syllable in some words while English does not and that, for learners to correctly apply in their writing the Spanish orthographic rules pertaining to stress, phonological awareness of stress is necessary (Beaudrie, 2007). A number of empirical studies have shown that L1 English L2 Spanish learners seem to have trouble identifying the stressed syllable in multisyllabic words and nonwords (Beaudrie, 2007; Romanelli & Menegotto, 2015; Romanelli et al., 2015). Thus, this population has been found to display low phonological awareness of stress. These findings led some to hypothesize that these L2 learners may have reduced perceptual sensitivity to the acoustic correlates of stress or that they may exploit them differently (in perceptual categorization) than L1 Spanish speakers, thus triggering the effect captured in the identification studies (Beaudrie, 2007; Ortega-Llebaria et al., 2013). Beaudrie (2007, p. 819) hypothesized that L1 English L2 Spanish learners have difficulty with their perception of the stressed syllable. And, indeed, Ortega-Llebaria et al. (2013) found that an orthogonal manipulation of several acoustic correlates of stress—pitch, duration, intensity—induced slight differences in the identification patterns of L1 and L2 Spanish speakers.

Ortín and Simonet (2022), however, pointed out that these studies (Beaudrie, 2007; Romanelli & Menegotto, 2015; Romanelli et al., 2015), due to the nature of their tasks, may have been able to, at most, document the existence of a reduced explicit phonological awareness of lexical stress in L1 English L2 Spanish learners but not (necessarily) an effect caused at the level of implicit phonological processing, let alone auditory perception. If one is to propose that the difficulties experienced by L2 learners are due to something occurring during auditory perception or processing, evidence of a different sort is needed—i.e., evidence at the level of implicit (not explicit) phonological processing and competence. For learners to develop phonological awareness of some feature, implicit phonological competence must indeed be in place, but implicit phonological competence may only be assessed with implicit phonological processing tasks, and not all difficulties with explicit phonological awareness derive from impoverished implicit phonological competence. Some effort has been dedicated to (i) verifying that L1 English

L2 Spanish learners indeed demonstrate reduced phonological competence with regards to stress when assessed with implicit tasks, and (ii) finding the *locus* of the perceptual difficulties these learners seem to encounter when processing Spanish stress auditorily (Kim, 2020; Ortega-Llebaria et al., 2013; Ortín & Simonet, 2022; Saalfeld, 2012; Sagarra & Casillas, 2018). Empirical findings have suggested that L2 Spanish learners whose L1 is English do experience some difficulties when perceiving stress contrasts, and that such difficulties could presumably have a phonological or perceptual basis, but these studies' tasks demanded very high cognitive loads or relied on lexical and grammatical competencies, thus obscuring the locus of the obstacle (Kim, 2020; Ortín & Simonet, 2022; Saalfeld, 2012; Sagarra & Casillas, 2018).

Our experiment was able to contribute new evidence on this issue. Firstly, our results were able to confirm that L1 English L2 Spanish learners demonstrate deadened perceptual processing of stress, which we investigated with a task that is able to implicitly address listeners' perceptual sensitivity to stress in phonological working memory without relying on lexical or grammatical knowledge (Kim, 2020; Ortega-Llebaria et al., 2013; Ortín & Simonet, 2022; Saalfeld, 2012). Secondly, our results suggest that the locus of the perceptual difficulty is not necessarily at the level of auditory perception, at least not principally, but at a higher level of processing, hence our referring to it as a difficulty found in phonological working memory or phonological processing, not auditory perception (Pisoni, 1973; Werker & Logan, 1985). As discussed, Ortega-Llebaria et al. (2013) found that L1 English L2 Spanish learners' reliance on suprasegmental acoustic cues differed from that of L1 Spanish speakers, thus suggesting an obstacle at the level of auditory perception. Recall, however, that the identification task in Ortega-Llebaria et al. (2013) may be said to tap into explicit phonological awareness rather than implicit processing. At any rate, we do not deny that differential acoustic cue weighting may contribute to the trouble L1 English L2 Spanish learners have with the processing of Spanish stress; we claim that there are other (additional) reasons for this difficulty. These reasons have to do with higher-level phonological processing and with phonological competence (or representation), not auditory perception *per se*.

In a study similar in purpose to the present one, Ortín and Simonet (2022) reported on the results of a sequence-recall experiment with a large sample of L1 English L2 Spanish learners. The sequence-recall task has been used to explore the phonological processing of stress of speakers of a variety of languages, including languages with and without contrastive lexical stress (Dupoux et al., 2001, 2010, 2008; Lin et al., 2014; Peperkamp et al., 2010; Qin et al., 2017). This task imposes high demands on working memory. Speakers of languages with contrastive stress, such as Spanish, are similarly accurate when reporting the order of sequences in trials whose minimal pairs differ as a function of stress than in baseline trials, in which tokens differ as a function of a phonemic segment. On the other hand, speakers of languages that lack contrastive stress, such as French, are more accurate in baseline trials than in trials that depend on stress contrasts. This has been interpreted as suggesting that speakers of languages that lack contrastive stress have

reduced, muted phonological processing of stress (relative to other phonological features and, secondarily, to speakers of languages with contrastive stress). The literature has also shown that stress-processing patterns vary in their strength, and that perceptual sensitivity to stress occurs on a spectrum, from hindered to heightened. The extremes seem to be occupied by speakers of languages that lack (French) or have (Spanish) contrastive lexical stress, whereas the middle is occupied by speakers of languages that lack contrastive stress but in which stress location is not entirely predictable from surface representations (Polish) (Peperkamp et al., 2010).

With a sequence-recall task, Ortín and Simonet (2022) found that the behavior of L1 English L2 Spanish learners also sits in the middle of the spectrum—for this group, accuracy was found to be lower in the stress condition than in the baseline condition, just like it had been found for speakers of French, Finnish, Hungarian, Korean, and Polish (Dupoux et al., 2001; Lin et al., 2014; Peperkamp et al., 2010; Qin et al., 2017). Ortín and Simonet (2022) proposed that, just like speakers of languages that lack contrastive stress can be in different places of the spectrum depending on whether stress is fully predictable (French) or not (Polish) in their language, speakers of languages that have contrastive stress may also occupy different locations on the spectrum depending on the functional load of stress in their language. Thus, L1 speakers of Spanish, in which stress has a relatively high functional load, demonstrate higher accuracy in a sequence-recall task (for stress contrasts) than L1 speakers of English, in which stress has a relatively low functional load; or, better, L1 English listeners have lower sensitivity to stress contrasts than to baseline contrasts while L1 Spanish speakers do not.

The research thus suggests that it is the lexical statistics of a feature, rather than merely the dichotomous classification of whether the feature is contrastive or not, that determine whether listeners demonstrate high(er) or low(er) perceptual sensitivity to the feature in implicit phonological working memory (Peperkamp et al., 2010). Such deadened perceptual sensitivity to stress is found in tasks that rely heavily on phonological working memory—i.e., tasks that demand high cognitive load—such as the sequence-recall task. Dupoux et al. (1997) found, however, that L1 French listeners were at ceiling in a simple perceptual discrimination task assessing their auditory categorization of stress patterns. Their auditory processing of stress was only negatively affected in tasks such as the ABX (Dupoux et al., 1997) and, above all, the sequence-recall (Dupoux et al., 2001). It seems that the difficulty found with the processing of stress has its locus, not (necessarily) in speech perception, but in phonological working memory (or phonological processing).

The present study contributes to the research on the implicit phonological processing of stress by evaluating L1 English L2 Spanish learners' perceptual sensitivity to stress in a task that taps into phonological working memory, like the sequence-recall task, but whose demands on cognitive load are much lighter than those of the sequence-recall task. Our findings suggest that L1 English L2 Spanish learners' deadened perceptual sensitivity to stress may be captured, not only in very

'difficult' tasks (Ortín & Simonet, 2022), but also in much 'easier' ones, such as the ABX. The ABX paradigm we used turned out to provide crucial data because it allowed us to capture a difference in the learners' behavior in response to adjacent and nonadjacent matching items. The difficulty with the processing of the matching items was circumscribed to the primacy-stress condition (ABA, BAB). Thus, in an ABX, identifying nonadjacent matching items was not difficult for the L2 learners, as long as the items in the trial did not differ solely on their stress configuration; and identifying matching items that differed solely on their stress configuration was not difficult, as long as the matching items were adjacent. This reveals an obstacle in phonological processing (or phonological working memory) rather than (only) acoustic memory or auditory perception (Pisoni, 1973; Werker & Logan, 1985).

In our experiment, the items in the triads had been recorded from three different talkers. Listeners, therefore, could not base their decisions exclusively on bottom-up, detailed acoustic information—abstraction was needed for them to identify differences and similarities in all triads. Nevertheless, the capacity of L1 English L2 Spanish learners to retain in working memory an abstract representation of stress seems to fade or dissipate rather rapidly—it is available only for a very short period or if no intervening auditory stimuli occur. These findings are reminiscent of a pattern found in Best et al. (2001) regarding the naïve, cross-linguistic perceptual discrimination of some Zulu consonants by L1 English speakers. Best et al. (2001) provides us with the theoretical framework or basis for us to interpret our findings.

In Best et al. (2001), a group of L1 English speakers with no knowledge of Zulu was asked to participate in an AXB discrimination task in which nonword minimal pairs differed as a function of some Zulu consonants that do not exist in English. The Perceptual Assimilation Model (PAM) (Best, 1995; Best & Tyler, 2007) postulates that the discriminability of contrasts not existing in a listeners' L1 depends on how the listener assimilates the members of the contrast (i.e., the sounds) to the sounds in their L1. The following are some of the cross-linguistic assimilation scenarios postulated by the PAM: (i) in single-category assimilation, the two sounds in the contrast are assimilated to the same sound in the L1 (in this case, discrimination is predicted to be poor); (ii) in two-category assimilation, the two sounds in the contrast are assimilated to two different sounds in the L1 (in this case, discrimination is expected to be optimal); (iii) in category-goodness assimilation, the two sounds in the contrast are assimilated to the same sound in the L1, but one is perceived to be more similar to the L1 sound than the other (in this case, discrimination is expected to be intermediate). There were three types of trials in Best et al. (2001), designed as a function of the types of cross-linguistic assimilations postulated by the PAM, but, most importantly, the study also analyzed any potential effects of recency. Since listeners presumably need to wait until they hear the third item for them to respond to an AXB triad, Best and colleagues hypothesized that accurately responding to primacy trials (AAB, BBA) would impose a heavier load on working-memory than responding to recency trials (ABB,

BAA)—the latency between exposure to the input and response is longer in primacy trials than in recency trials.

A relevant finding in Best et al. (2001) was that L1 English listeners demonstrated recency effects only for contrasts classified as single-category assimilation scenarios, not in the other types of contrasts: In single-category assimilation contrasts, listeners were less accurate in primacy trials (AAB, BBA) than in recency trials (ABB, BAA). Best and colleagues suggested that this finding derives from the information listeners rely on to make their decision in each trial. On the one hand, in two-category assimilation, listeners can use phonetic or phonological (that is, linguistic) information to respond, as each item is assimilated to a different abstract category available in the listeners' mental representation and it is those categories listeners rely on. On the other hand, in single-category assimilation, listeners may not rely on linguistic information to respond, since the two members of the minimal pair are assimilated to the same L1 abstract category; in fact, relying on linguistic information would absolutely hinder discrimination in this scenario. For listeners to be accurate in their discrimination of single-category assimilation contrasts, their discrimination must be based on nonlinguistic, acoustic (or auditory) information. It is known that, in working memory, acoustic detail (auditory memory) is much more vulnerable to the effects of time (latency) than abstract linguistic (phonetic, phonological) information (Pisoni, 1973; Pisoni & Tash, 1974; Werker & Logan, 1985), and this could explain the asymmetry found in Best et al. (2001). Thus, listeners, when trying to discriminate two sounds that they assimilate to a single L1 category are forced to rely on auditory memory; however, since this type of memory rapidly dissipates from working memory, perceptual sensitivity to the contrast is higher when stimuli are adjacent, or latency is short, than when stimuli are not adjacent, or latency is long. We surmise that the same type of explanation may be used to account for the findings of the present study.

We postulate that L1 English speakers (and L1 English L2 Spanish learners) do not possess a robust, detailed mental representation of lexical stress independent of segmental representation (Cutler, 1986). This does not mean that they do not 'hear' suprasegmental acoustic cues during speech comprehension. In fact, L1 English listeners must be sensitive to suprasegmental cues (Chrabaszcz, Winn, Lin, & Idsardi, 2014; Fry, 1955, 1958; Ortega-Llebaria et al., 2013), among other things because such cues may be sporadically used to distinguish lexical minimal pairs in their L1. However, the low functional load of lexical stress in English likely provides insufficient motivation or input for L1 English speakers to ever develop strong, robust abstract representations of lexical stress (Cooper et al., 2002; Cutler, 1986, 2005; Warner & Cutler, 2017). Thus, when L1 English speakers are asked to participate in a perceptual discrimination experiment, they may be able to rely on relatively detailed auditory or acoustic (that is, nonlinguistic) memory, with which they have some practice, but not on abstract linguistic (that is, phonetic or phonological) representations. Since, in working memory, latency affects auditory memory much more rapidly

than it affects linguistic memory, discrimination scenarios that rely heavily on acoustic rather than linguistic memory will yield latency effects (in the form of a recency effect). L1 Spanish speakers, on the other hand, possess a robust, detailed mental representation of lexical stress (Dupoux et al., 1997, 2001; Ortín & Simonet, 2022; Peperkamp et al., 2010; Soto-Faraco et al., 2001), and they can rely on it in perceptual discrimination experiments. Since, in working memory, linguistic representations are relatively impervious to recency effects, L1 Spanish speakers are less likely than L1 English speakers to manifest any effects of recency. Our interpretation of the recency effects in our findings lead us to postulate that L1 English speakers (and L1 English L2 Spanish learners) lack a phonological representation of stress and thus cannot rely on it in phonological processing.

Lastly, let us discuss the findings pertaining to L2 experience. The present study did not find any effects of L2 experience or proficiency on the perceptual processing of stress; there was no evidence that L1 English L2 Spanish learners become more attuned to the processing of lexical stress with increased experience with Spanish. This finding is in line with those pertaining to L1 French L2 Spanish learners (Dupoux et al., 2010, 2008) as well as those in Saalfeld (2012), a study that included a sample of relatively inexperienced L1 English L2 Spanish learners. Other studies with samples of L1 English L2 Spanish learners, however, have produced positive, albeit modest, effects of experience (Ortín & Simonet, 2022; Sagarra & Casillas, 2018). In Sagarra and Casillas (2018), experienced L2 learners of Spanish, but not inexperienced ones, were found to exploit lexical stress to resolve lexical competition during spoken word recognition. Ortín and Simonet (2022) found that an increase in proficiency—measured by means of lexical and grammatical knowledge tests—was positively, albeit very modestly, associated with an increase in accuracy in a sequence-recall task assessing the processing of stress. The sample in Ortín and Simonet (2022) is very similar, in terms of their experience and proficiency, to the sample of the current study, whereas the experienced L2 learners in Sagarra and Casillas (2018) were much more experienced than the most experienced participants in our current sample. A comparison of the findings of these studies with those of the current one suggests that intermediate-level learners of Spanish may improve in their phonological processing of Spanish stress with increased experience in the language, but progress may be shown first in highly demanding or sensitive tasks (the sequence-recall task) than in less demanding or sensitive tasks (AXB). Only very experienced learners, however, may exploit stress contrasts during online lexical searches and may develop a higher sensitivity to stress (Sagarra & Casillas, 2018).

5. Conclusion

A total of 87 people participated in an ABX categorical matching task with triads of auditory stimuli minimally contrasting in stress (target) or segmental composition (baseline). Matching items in the triads could be adjacent (recency condition) or not adjacent (primacy condition).

Auditory stimuli were recorded from L1 Spanish talkers, and they consisted exclusively of nonwords. Participants were divided into two groups: a small control group comprised of L1 Spanish speakers, and a large group of L1 English L2 Spanish learners of varying linguistic proficiencies in their L2. The most significant results concerned the L1 English L2 Spanish learners. Unlike the L1 Spanish controls, the L2 learners were less accurate in the primacy than in the recency condition, and they were also less accurate in the stress than in the baseline condition. With regards to accuracy, when examining the interaction of the two effects, it was found that the L2 learners were at ceiling in three types of trials (both the recency and primacy trials in the baseline condition, and at the intersection of recency and stress). The learners' perceptual sensitivity was compromised only at the intersection of primacy and stress, that is, in trials in which the matching stimuli were not adjacent. The pattern in the findings suggest that L1 English L2 Spanish learners' acquisitional difficulties with Spanish stress are likely to be due to a reduced perceptual sensitivity to stress distinctions that manifests itself in working memory (not acoustic, auditory memory) (Dupoux et al., 1997, 2001; Peperkamp et al., 2010; Pisoni, 1973; Werker & Logan, 1985). We conclude that the compromised perceptual sensitivity to stress contrasts shown by this population is likely due to a null or 'blurry' representation of stress as a phonological category in their mental grammar—with cascading effects for phonological processing—and this is likely inherited from these learners' L1, English.

Additional file

The additional file for this article can be found as follows:

- **Supplementary Materials.** Design of ABX task: items, trials, and factor levels. DOI: <https://doi.org/10.16995/labphon.7978.s1>

Competing interests

The authors have no competing interests to declare.

References

- Beaudrie, S. (2007). La adquisición del acento ortográfico en la clase de español como segunda lengua. *Hispania*, 90, 809–823. DOI: <https://doi.org/10.2307/20063614>
- Beaudrie, S. (2017). The teaching and learning of spelling in the Spanish heritage language classroom: Mastering written accent marks. *Hispania*, 100, 596–611. <https://www.jstor.org/stable/26387811>. DOI: <https://doi.org/10.1353/hpn.2017.0101>
- Best, C. T. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-language Research* (pp. 171–204). York Press.
- Best, C. T., McRoberts, G. W., & Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *The Journal of the Acoustical Society of America*, 109, 775–794. DOI: <https://doi.org/10.1121/1.1332378>
- Best, C. T., McRoberts, G. W., & Sithole, N. M. (1988). Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology. Human Perception and Performance*, 14, 345–360. DOI: <https://doi.org/10.1037/0096-1523.14.3.345>
- Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In O.-S. Bohn & M. Munro (Eds.), *Language experience in second language speech learning: In honor of James Emil Flege* (pp. 13–34). John Benjamins. DOI: <https://doi.org/10.1075/llt.17.07bes>
- Calin-Jageman, R., & Cumming, G. (2019). The New Statistics for Better Science: Ask How Much, How Uncertain, and What Else is Known. *The American Statistician*, 73, 271–280. DOI: <https://doi.org/10.1080/00031305.2018.1518266>
- Carney, A. E., Widin, G. P., & Viemeister, N. F. (1977). Noncategorical perception of stop consonants differing in VOT. *The Journal of the Acoustical Society of America*, 62, 961–970. DOI: <https://doi.org/10.1121/1.381590>
- Chen, M. (1970). Vowel length variation as a function of the voicing of the consonant environment. *Phonetica*, 22, 129–159. DOI: <https://doi.org/10.1159/000259312>
- 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: JSLHR*, 57, 1468–1479. DOI: https://doi.org/10.1044/2014_JSLHR-L-13-0279

- Colantoni, L., Steele, J., & Escudero, P. (2015). *Second Language Speech: Theory and Practice*. Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9781139087636>
- 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, 207–228. DOI: <https://doi.org/10.1177/00238309020450030101>
- Cumming, G. (2013). *Understanding the New Statistics: Effect Sizes, Confidence Intervals, and Meta-Analysis*. Routledge. DOI: <https://doi.org/10.4324/9780203807002>
- Cumming, G., & Calin-Jageman, R. (2016). *Introduction to the new statistics: Estimation, open science, and beyond*. Routledge. DOI: <https://doi.org/10.4324/9781315708607>
- Cutler, A. (1986). Forbear is a homophone: Lexical prosody does not constrain lexical access. *Language and Speech*, 29, 201–220. DOI: <https://doi.org/10.1177/002383098602900302>
- Cutler, A. (2005). Lexical stress. In D. Pisoni & R. Remez (Eds.), *The Handbook of Speech Perception* (pp. 264–289). Blackwell Publishing Ltd. DOI: <https://doi.org/10.1002/9780470757024.ch11>
- Cutler, A. (2012). *Native Listening: Language Experience and the Recognition of Spoken Words*. MIT Press. DOI: <https://doi.org/10.7551/mitpress/9012.001.0001>
- Cutler, A., Norris, D., & Sebastián-Gallés, N. (2004). Phonemic repertoire and similarity within the vocabulary. In S. Kin & M. J. Bae (Eds.), *Interspeech 2004*. Sunjijn Printing Co. DOI: <https://doi.org/10.21437/Interspeech.2004-61>
- Cutler, A., & Pasveer, D. (2006). Explaining cross-linguistic differences in effects of lexical stress on spoken-word recognition. In R. Hoffman & H. Mixdorff (Eds.), *Speech Prosody 2006*. TUD Press.
- Dahan, D., & Magnuson, J. S. (2006). Spoken word recognition. In M. J. Traxler & M. A. Gernsbacher (Eds.), *Handbook of Psycholinguistics (Second Edition)* (pp. 249–283). Academic Press. DOI: <https://doi.org/10.1016/B978-012369374-7/50009-2>
- Diehl, R. L., Lotto, A. J., & Holt, L. L. (2004). Speech perception. *Annual Review of Psychology*, 55, 149–179. DOI: <https://doi.org/10.1146/annurev.psych.55.090902.142028>
- Domahs, U., Knaus, J., Orzechowska, P., & Wiese, R. (2012). Stress “deafness” in a language with fixed word stress: An ERP study on Polish. *Frontiers in Psychology*, 3, 439. DOI: <https://doi.org/10.3389/fpsyg.2012.00439>
- Dupoux, E., Pallier, C., Sebastián-Gallés, N., & Mehler, J. (1997). A destressing “deafness” in French? *Journal of Memory and Language*, 36, 406–421. DOI: <https://doi.org/10.1006/jmla.1996.2500>
- Dupoux, E., Peperkamp, S., & Sebastián-Gallés, N. (2001). A robust method to study stress “deafness.” *The Journal of the Acoustical Society of America*, 110, 1606–1618. DOI: <https://doi.org/10.1121/1.1380437>
- Dupoux, E., Peperkamp, S., & Sebastián-Gallés, N. (2010). Limits on bilingualism revisited: Stress “deafness” in simultaneous French-Spanish bilinguals. *Cognition*, 114, 266–275. DOI: <https://doi.org/10.1016/j.cognition.2009.10.001>

- Dupoux, E., Sebastián-Gallés, N., Navarrete, E., & Peperkamp, S. (2008). Persistent stress “deafness”: The case of French learners of Spanish. *Cognition*, *106*, 682–706. DOI: <https://doi.org/10.1016/j.cognition.2007.04.001>
- Escudero, P. (2005). *Linguistic perception and second language acquisition: Explaining the attainment of optimal phonological categorization*. LOT Dissertation Series 113.
- Face, T. L. (2005). Syllable weight and the perception of Spanish stress placement by second language learners. *Language Learning and Development: The Official Journal of the Society for Language Development*, *3*, 90–103.
- Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 229–273). York Press.
- Flege, J. E., Aoyama, K., & Bohn, O.-S. (2021). The revised Speech Learning Model (SLM-r) applied. In R. Wayland (Ed.), *Second Language Speech Learning: Theoretical and Empirical Progress* (pp. 84–118). Cambridge University Press. DOI: <https://doi.org/10.1017/9781108886901.003>
- Flege, J. E., & Bohn, O.-S. (2021). The revised Speech Learning Model (SLM-r). In R. Wayland (Ed.), *Second Language Speech Learning: Theoretical and Empirical Progress* (pp. 3–83). Cambridge University Press. DOI: <https://doi.org/10.1017/9781108886901.002>
- Fry, D. B. (1955). Duration and intensity as physical correlates of linguistic stress. *The Journal of the Acoustical Society of America*, *27*, 765–768. DOI: <https://doi.org/10.1121/1.1917773>
- Fry, D. B. (1958). Experiments in the perception of stress. *Language and Speech*, *1*, 126–152. DOI: <https://doi.org/10.1177/002383095800100207>
- Gertken, L. M., Amengual, M., & Birdsong, D. (2014). Assessing language dominance with the Bilingual Language Dominance Profile. In P. Leclercq, A. Edmonds, & H. Hilton (Eds.), *Measuring L2 Proficiency: Perspectives from SLA* (pp. 208–225). Bilingual Matters. DOI: <https://doi.org/10.21832/9781783092291-014>
- Giegerich, H. J. (1992). *English Phonology: An Introduction*. Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9781139166126>
- Gillon, G. (2017). *Phonological Awareness: From Research to Practice (2nd edition)*. Guilford Press.
- Hualde, J. I. (2005). *The Sounds of Spanish*. Cambridge University Press.
- Izura, C., Cuetos, F., & Brysbaert, M. (2014). Lextale-Esp: A test to rapidly and efficiently assess the Spanish vocabulary size. *Psicologica: International Journal of Methodology and Experimental Psychology*, *35*, 49–66.
- Jessen, M., & Ringen, C. (2002). Laryngeal features in German. *Phonology*, *19*, 189–218. DOI: <https://doi.org/10.1017/S0952675702004311>
- Kim, J. Y. (2020). Discrepancy between heritage speakers’ use of suprasegmental cues in the perception and production of Spanish lexical stress. *Bilingualism: Language and Cognition*, *23*, 233–250. DOI: <https://doi.org/10.1017/S1366728918001220>
- Lenth, R. (2018). *emmeans: Estimated marginal means, aka Least-squares means* (1.4.7) [Computer software]. <https://cran.r-project.org/package-emmeans>

- Lin, C. Y., Wang, M., Idsardi, W. J., & Xu, Y. (2014). Stress processing in Mandarin and Korean second language learners of English. *Bilingualism: Language and Cognition*, 17, 316–346. DOI: <https://doi.org/10.1017/S1366728913000333>
- Mann, V. A. (1986). Distinguishing universal and language-dependent levels of speech perception: Evidence from Japanese listeners' perception of English "l" and "r." *Cognition*, 24(3), 169–196. DOI: [https://doi.org/10.1016/S0010-0277\(86\)80001-4](https://doi.org/10.1016/S0010-0277(86)80001-4)
- Martínez García, M. T. (2016). *Tracking bilingual activation in the processing and production of Spanish stress* (Doctoral dissertation). University of Kansas.
- Morrill, T. (2012). Acoustic correlates of stress in English adjective-noun compounds. *Language and Speech*, 55, 167–201. DOI: <https://doi.org/10.1177/0023830911417251>
- Nadeu, M. (2014). Stress- and speech rate-induced vowel quality variation in Catalan and Spanish. *Journal of Phonetics*, 46, 1–22. DOI: <https://doi.org/10.1016/j.wocn.2014.05.003>
- 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. DOI: <https://doi.org/10.1016/j.wocn.2013.01.006>
- Ortega-Llebaria, M., & Prieto, P. (2011). Acoustic correlates of stress in Central Catalan and Castilian Spanish. *Language and Speech*, 54, 73–97. DOI: <https://doi.org/10.1177/0023830910388014>
- Ortín, R., & Simonet, M. (2022). Phonological processing of stress by native English speakers learning Spanish as a second language. *Studies in Second Language Acquisition*, 44, 460–482. DOI: <https://doi.org/10.1017/S0272263121000309>
- Peirce, J. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods*, 162, 8–13. DOI: <https://doi.org/10.1016/j.jneumeth.2006.11.017>
- Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, 51, 195–203. DOI: <https://doi.org/10.3758/s13428-018-01193-y>
- Peperkamp, S., Vendelin, I., & Dupoux, E. (2010). Perception of predictable stress: A cross-linguistic investigation. *Journal of Phonetics*, 38, 422–430. DOI: <https://doi.org/10.1016/j.wocn.2010.04.001>
- Pisoni, D. (1973). Auditory and phonetic memory codes in the discrimination of consonants and vowels. *Perception & Psychophysics*, 13, 253–260. DOI: <https://doi.org/10.3758/BF03214136>
- Pisoni, D., & Tash, J. (1974). Reaction times to comparisons within and across phonetic categories. *Perception & Psychophysics*, 15, 285–290. DOI: <https://doi.org/10.3758/BF03213946>
- Qin, Z., Chien, Y.-F., & Tremblay, A. (2017). Processing of word-level stress by Mandarin-speaking second language learners of English. *Applied Psycholinguistics*, 38, 541–570. DOI: <https://doi.org/10.1017/S0142716416000321>
- Romanelli, S., & Menegotto, A. C. (2015). English speakers learning Spanish: Perception issues regarding vowels and stress. *Journal of Language Teaching and Research*, 6, 30–42. DOI: <https://doi.org/10.17507/jltr.0601.04>

- Romanelli, S., Menegotto, A. C., & Smyth, R. (2015). Stress perception: Effects of training and a study abroad program for L1 English late learners of Spanish. *Journal of Second Language Pronunciation*, 1, 181–210. DOI: <https://doi.org/10.1075/jslp.1.2.03rom>
- Saalfeld, A. K. (2012). Teaching L2 Spanish Stress. *Foreign Language Annals*, 45, 283–303. DOI: <https://doi.org/10.1111/j.1944-9720.2012.01191.x>
- Sagarra, N., & Casillas, J. V. (2018). Suprasegmental information cues morphological anticipation during L1/L2 lexical access. *Journal of Second Language Studies*, 1, 31–59. DOI: <https://doi.org/10.1075/jsls.17026.sag>
- Samuel, A. G. (2011). Speech perception. *Annual Review of Psychology*, 62, 49–72. DOI: <https://doi.org/10.1146/annurev.psych.121208.131643>
- Samuel, A. G. (2020). Psycholinguists should resist the allure of linguistic units as perceptual units. *Journal of Memory and Language*, 111, 104070. DOI: <https://doi.org/10.1016/j.jml.2019.104070>
- Simonet, M. (2016). The phonetics and phonology of bilingualism. In *Oxford Handbooks Online: Scholarly Research Reviews* (pp. 1–25). Oxford University Press. DOI: <https://doi.org/10.1093/oxfordhb/9780199935345.013.72>
- Singman, H., Bolker, B., Westfall, J., Aust, F., & Ben-Shachar, M. S. (2020). *afex: Analysis of factorial experiments* (0.27–2) [Computer software]. <https://cran.r-project.org/package=afex>
- Sluijter, A. M. C., & van Heuven, V. J. (1996a). Spectral balance as an acoustic correlate of linguistic stress. *The Journal of the Acoustical Society of America*, 100, 2471–2485. DOI: <https://doi.org/10.1121/1.417955>
- Sluijter, A. M. C., & van Heuven, V. J. (1996b). Acoustic correlates of linguistic stress and accent in Dutch and American English. *Proceedings of 4th International Conference on Spoken Language Processing*, 2, 630–633. DOI: <https://doi.org/10.1109/ICSLP.1996.607440>
- Smith, B. L., Hayes-Harb, R., Bruss, M., & Harker, A. (2009). Production and perception of voicing and devoicing in similar German and English word pairs by native speakers of German. *Journal of Phonetics*, 37, 257–275. DOI: <https://doi.org/10.1016/j.wocn.2009.03.001>
- Soto-Faraco, S., Sebastián-Gallés, N., & Cutler, A. (2001). Segmental and suprasegmental mismatch in lexical access. *Journal of Memory and Language*, 45, 412–432. DOI: <https://doi.org/10.1006/jmla.2000.2783>
- The Jamovi Project. (2020). *jamovi* (1.2.22) [Computer software]. <https://www.jamovi.org>
- Torreira, F., Simonet, M., & Hualde, J. I. (2014). Quasi-neutralization of stress contrasts in Spanish. In N. Campbell, D. Gibbon, & D. J. Hirst (Eds.), *Speech Prosody 7* (pp. 197–201). Trinity College, Dublin. DOI: <https://doi.org/10.21437/SpeechProsody.2014-27>
- van Donselaar, W., Koster, M., & Cutler, A. (2005). Exploring the role of lexical stress in lexical recognition. *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology*, 58, 251–273. DOI: <https://doi.org/10.1080/02724980343000927>
- van Leussen, J.-W., & Escudero, P. (2015). Learning to perceive and recognize a second language: The L2LP model revised. *Frontiers in Psychology*, 6, 1–12. DOI: <https://doi.org/10.3389/fpsyg.2015.01000>

Warner, N., & Cutler, A. (2017). Stress effects in vowel perception as a function of language-specific vocabulary patterns. *Phonetica*, *74*, 81–106. DOI: <https://doi.org/10.1159/000447428>

Werker, J. F., & Logan, J. S. (1985). Cross-language evidence for three factors in speech perception. *Perception & Psychophysics*, *37*, 35–44. DOI: <https://doi.org/10.3758/BF03207136>

Wickham, H., Averick, M., Bryan, J., Chang, W., D'Agostino McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Lin Pedersen, T., Miller, E., Milton Bache, S., Müller, K., Ooms, J., Robinson, D., Paige Seidel, D., Spinu, V., ... Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*, 1686. DOI: <https://doi.org/10.21105/joss.01686>

