Use of Word-Level Stress in L2 Spanish Word Recognition

María Teresa Martínez García[1], *Julie Amber*[2], *Sandra Schwab*[2]

[1]University of Utah, Asia Campus
[2]University of Bern

maria.martinezgarcia@utah.edu, julie.kamber@unibe.ch, sandra.schwab@unibe.ch

Abstract

Individuals with different L1s rely differently on suprasegmental cues when recognizing spoken words in L1 and in L2. The present study investigates the use of stress information in L2 word recognition by intermediate-level learners of Spanish with different L1s: German (i.e., a language with distinctive word stress; DE), French and Korean (i.e., languages without distinctive word stress; FR, KR). Contrary to DE, the absence of word stress in French and Korean was expected to hinder FR and KR ability to use stress in L2 Spanish word recognition. In a cross-modal word-identification task, participants listened to semantically ambiguous auditory sentences ending with incomplete two-syllable word fragments and had to choose the word that matched the heard fragment, which either contained (stressed condition) or lacked a stressed syllable (unstressed condition). While the results revealed no difference between three language groups in the unstressed condition (62%), the groups differed in the stressed condition. KR accuracy (77%) was unexpectedly higher than FR (63%) and as good as DE (76%), suggesting that, contrary to FR, KR were able to use stress to access L2 words, although word stress does not exist in their L1.

Index Terms: speech perception, word recognition, lexical stress, Spanish, French, Korean, German

1. Introduction

1.1. Lexical stress as a cue for L2 word recognition

Listeners from different language backgrounds do not equally rely on suprasegmental information during online word recognition. For example, since Spanish has lexical stress and does not have vowel reduction, suprasegmental information in Spanish can distinguish among competing lexical hypotheses, not only for minimal pairs (e.g., *PApa* 'potato' vs. *paPÁ* 'father'), but also for temporarily overlapping pairs (e.g., *peLOta* 'ball' vs. *peloTÓN* 'platoon') [1].

Many scholars frequently associate challenges in assigning stress in a second language (L2) with the influence of an individual's native language (L1). In essence, individuals whose first languages (L1) utilize stress as a means of distinguishing meaning (e.g., Spanish, English, German) tend to exhibit a greater aptitude in perceiving stress compared to those whose native languages (e.g., French, Finnish, Korean) do not depend on stress for conveying meaning ([2]), phenomenon known as "stress deafness" ([2] [3] [4]). For instance, French listeners were less accurate than Spanish listeners in perceiving lexical stress in Dutch ([2]) and had difficulties with L2 Spanish stress ([3]). However, individual differences ([5]) and task effects ([6]) could also explain some of these results.

Accordingly, while native speakers rely on word-level stress for word recognition, nonnative speakers whose L1 uses lexical stress can make use of these cues for recognizing words in their L2 ([6] [7]). For example, in the presence of the suprasegmental cues that signal word-level stress in Spanish (positive cue), both native speakers and English-speaking learners of Spanish can rely on this cue for word recognition. On the other hand, when the suprasegmental cues are not present to cue word-level stress (negative cue), both native speakers and learners struggle to use the absence of cues for word recognition ([8]). However, studies looking at how native speakers whose L1 does not have lexical stress use this cue for word recognition in their L2 are inconclusive, with some studies supporting the claim that these individuals are 'deaf' to this feature ([3] [4] [6]), while other studies seem to indicate that, at least some individuals, can use lexical stress as a cue for L2 word recognition ([9] [10]).

Finally, most of the previous research has focused on English-speaking learners of Spanish and how they manage to employ lexical stress in Spanish despite the absence of segmental cues (no vowel reduction) [3,4], yet other language combinations are still not well-studied. The present research will add to the body of literature on this field by testing the use of word-level stress in L2 Spanish by native speakers of German, French, and Korean, because of the differences in the stress properties among these languages.

1.2. Word-level stress in Spanish, German, French, and Korean

The languages examined in this study—Spanish, German, French, and Korean—vary in their accentual characteristics.

Spanish is a free-stress language, where lexical stress serves a distinct role ([5]), differentiating words that are segmentally identical, such as *número* (/'numero/, "(the) number") and *numero* (/nu'mero/, "I number"). In Spanish, word-level stress is signaled with suprasegmental cues (F0, duration, and intensity). Stress placement can be predicted by abstract, complex stress assignment rules: For nouns, the so-known regular stress rule states that stress falls on the last syllable if it ends with a consonant other than [n] or [s], and otherwise on the penultimate syllable ([22]). Like in the case of Spanish, lexical stress in German is complex, rule-governed, and highly predictable. However, and contrary to Spanish, lexical stress in German employes both suprasegmental (F0, duration, and amplitude) and segmental (vowel reduction) cues ([6,7]).

In contrast, French and Korean are languages without word stress. In the case of French, it is a language with fixed stress, occurring on the final syllable of the Accentual Phrase (AP) ([11]). Consequently, stress acts as a boundary marker of each accentual phrase (i.e., it signifies the end of the accentual phrase; [11]). Furthermore, French shows what is known as

"syncretism" between accentuation and intonation ([12]), that is, the realization of primary stress and the contours of intonation are centered on the ultimate syllable of the AP, resulting in a blending of accentual and intonational structures. Similarly, Korean has also been described as a language without word-level stress (i.e., [10]), which utilizes prominence at the phrase level, also positioned at the conclusion of the AP, primarily indicated by f0 to denote phrasal boundaries. In standard (Seoul) Korean, the AP is marked by a high tone at its end, while the initial boundary of the prosodic unit could feature either a high or a low tone, depending on whether the first element of the utterance is lenis, aspirated, or fortis (e.g., [13] [14]). Although certain Korean dialects (e.g., Chonnam or North Kyungsang) has been described as maintaining certain distinctions (e.g., vowel length, as in *sa:.kwa* 'apology' vs. *sa.kwa* 'apple') ([15] [16] [17]), Seoul Korean (the dialect under examination) has discarded this characteristic.

The current study aims to investigate whether German, French, and Korean listeners at an intermediate-to-advanced proficiency level in L2 Spanish would make use of stress for word recognition in Spanish in the presence of stress (a 'positive' cue), in comparison with a condition where there is absence of stress (a 'negative' cue) and how this feature may constrain word recognition in all groups ([17]).

2. Methodology

2.1. Participants

Three groups completed the study. The first group included 20 Swiss German listeners (hereafter 'DE') recruited at Universities of Bern and Zurich, Switzerland (mean age: 23.7 years, sd: 2.29, age range: 19-27 years). The second group included Swiss French listeners (hereafter 'FR') recruited at the University of Fribourg, Switzerland (mean age: 22.9 years, sd: 2.75, age range: 19-29 years). The third group included 20 Korean listeners (hereafter 'KR') recruited at the University of Utah Asia Campus and at Hankuk University of Foreign Studies, South Korea (mean age: 23.8 years, sd: 2.8, age range: 19-30 years).

Based on the Spanish version of Dialang listening proficiency test ([19], scale from 1-beginner to 6-advanced), the three groups were intermediate learners of Spanish (median = 3) although with different degree of within-group variability (DE: IQR = 1, range = 2-6; FR: IQR = 1, range = 1-5; KR: IQR = 1.25, range = 2-6). None of the participants started learning either Spanish nor English before the age of 9 years old, and their English proficiency ranged from beginner to advanced (tested with the English version of the Dialang listening proficiency test). Participants in the three groups had similar experiences learning Spanish. The DE group on average had studied Spanish for 4.24 years (sd: 1.9, range: 2-8 years), the FR group on average 3.67 years (sd: 2.2, range: 1-8 years), and the KR group on average 5 years (sd: 2.8, range: 2-12 years). Finally, the DE group is the one with the longest experience in a Spanish-speaking country (average: 9.32 months, sd: 13.8), following by the KR group (average: 6.41 months, sd: 9.9) and the FR group (average: 2.66 months, sd: 5.3).

2.2. Procedure

The experiment was administered using Paradigm ([20]). Participants completed a cross-modal word-identification task (adapted from [6] [7] [21]). They heard semantically ambiguous auditory sentences that ended with two-syllable word fragments (e.g., *Elena dijo peLO…* 'Elena said peLO…') presented in two possible conditions. In the Stressed condition, the fragments were stressed on the penultimate syllable (e.g., peLO-, from the word *peLOta* 'ball'); in the Unstressed condition, the fragments were unstressed (e.g., pelo-, from the word *peloTÓN* 'squad'). Participants were asked to choose the word corresponding to the fragment they heard by clicking on one of the two options that appeared on the screen (e.g., "peLOta" vs. "peloTÓN").

Participants were also asked to rate their familiarity with the items that appeared in the word recognition task, using a 5-point scale, where 5 points indicated the ability to understand and properly use the word in a sentence. The word familiarity scores were used as an objective measure to assess participants' proficiency of Spanish in order to take into account the interindividual differences regarding Spanish proficiency. Indeed, although the three groups' proficiency level and L2 learning experience were globally similar, their degree of variability within each group (i.e., interindividual variability) differed. Thus, in addition to measuring to what extent the participants knew the words included in the experiment, the word familiarity scores, viewed as an objective proxy for L2 proficiency, allowed us to account for this interindividual variability.

2.3. Stimuli

Twenty-four experimental stimuli were created and distributed into four lists in a counterbalanced way. All of the experimental items were trisyllabic words with 'regular' stress placement ([22]). Words with penultimate stress (Stressed fragment) and final stress (Unstressed fragments) were matched in terms of lexical frequency. All fragments belonged to words that follow regular stress patterns in Spanish and no visual information regarding the stress placement was used, other than the regular diacritics. An example of the two conditions can be seen in Table 1.

Table 1: *Example of an experimental trial in the two stress conditions (underlined is the correct response).*

AUDIO	WORD CHOICE	
Stressed Fragment		
peLO-	pelota	pelotón
Unstressed Fragment		
pelo-	pelota	pelotón

The experimental task also included 36 filler items, in which the difference between the two potential target words was segmental rather than suprasegmental (e.g., *balido* 'bleat' vs. *batido* 'shake'). Both the experiment trials and the filler items were included in a randomized order in the word familiarity task.

2.4. Data analysis

Statistical analyses were carried out using R (version 4.0.3; R Development Core Team, 2022; lme4 R package; [24]). A mixed-effects logistic regression with the correct/incorrect responses was modelled as binary output ([25]). The fixed part comprised 'L1' (DE, FR, KR), 'Stress' (Unstressed, Stressed), and the interaction between them. The nominal variables 'L1' and 'Stress' were recoded into [0, 1] dummy variables. Word familiarity (centered to the mean) was included as a control

variable. The random part of the model included random intercepts for participants and items as well as random slope allowing the capture of the differential effect of 'Stress' across participants. The significance of the main and interaction effects was assessed with likelihood ratio tests that compared the model with the main or interaction effect to a model without such effect. The estimates (β), expressed in logit, were computed taking 'incorrect response' as the reference level for the dependent variable (i.e., correct/incorrect responses). The figures in the results section show percentage of correct responses, although the models were run on raw data (correct/incorrect responses).

3. Results

Figure 1 shows the results for the DE, FR, and KR groups in the Unstressed (left) and Stressed (right) conditions, respectively.

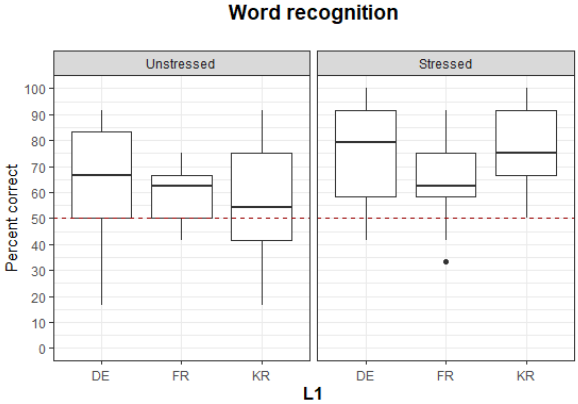

Figure 1: *Percent correct by L1 (DE, FR, KR) in the Unstressed (left) and the Stressed (right) conditions.*

Results showed no main effect of L1 ($\chi^2(2) = 0.85$, p = .65), but a marginal effect of stress ($\chi^2(1) = 3.04$, p = .08), suggesting that there was a difference between the Stressed and Unstressed conditions. Importantly, results showed an interaction between Stress and L1 ($\chi^2(2) = 6.32$, p = .042). Post-hoc analysis (with Tukey correction) revealed than when in the Unstressed condition, the performance of the three language groups did not differentiate from each other (p = [.06 - .93]). In the Stressed condition, KR unexpectedly did not differ from DE (β= 0.233, SE = 0.321, z = 0.727, p = .747), but presented a significantly higher performance than FR (β= 0.874, SE = 0.304, z = 2.870, p = .011). Surprisingly, DE marginally differ from FR (β= 0.640, SE = 0.307, z = 2.083, p = .093).

Having a look at the difference between the Unstressed and Stressed conditions for each language group, we observe that KR present a significantly better performance in Stressed than in Unstressed condition (β= 1.092, SE = 0.344, z = 3.178, p = .002), whereas the difference in DE goes in the same direction but just fail to reach significance (β= 0.604, SE = 0.345, z = 1.750, p = .080). In contrast, FR performance does not differ in both conditions (β= 0. 122, SE = 0.330, z = 0.369, p = .712).

4. Discussion

The current study aimed to explore whether German, French, and Korean listeners at an intermediate-to-advanced proficiency level in L2 Spanish could make use of stress for word recognition in Spanish and how the properties of the native language (free stress, as in the case of German, or fixed stress, as in the cases of French and Korean) may modify this effect. The results regarding the Stressed condition indicate that, as predicted, native speakers of a language with lexical stress (DE group) present a high performance, suggesting that they were able to make use of this cue when present in the acoustic input. Also as predicted, FR listeners presented a lower performance, suggesting that they could not make use of lexical stress in L2 Spanish for word recognition. Surprisingly, the KR group showed that, although their native language does not have lexical stress, they were able to use stress for word recognition in L2 Spanish, as indicated by a higher accuracy in the word recognition task.

The DE and KR groups showed a similar pattern: The overall accuracy is higher in the Stressed fragment condition than Unstressed condition (once again, the FR group did not show any statistical difference in performance between the two conditions). That is, when there is a positive cue in the acoustic input, it helps access the correct lexical item, while negative cues (absence of such information) do not have such a strong facilitatory effect, at least for the DE and KR listeners. Thus, these L2 learners appear to rely on similar mechanisms as native speakers for recognizing Spanish words, as the same pattern has been attested before in the literature ([8]).

The results of the DE and KR groups are in line with previous findings showing that L2 learners can show sensitivity to stress cues for word recognition (for studies on English learners of Spanish, see: [27]; for studies on Korean learners of English, see: [10] [28]). On the other hand, the results of the FR group seem to indicate the persistence of 'stress deafness' among those individuals whose native language does not have lexical stress ([2] [3] [4]) is not only persistent, but it also transfers to word recognition, making it more challenging for this individual to successfully recognize words online.

While FR and KR were predicted to pattern similarly based on the fact that neither language has lexical stress, and they were expected to pattern poorly in the word recognition task (compared with the DE group). However, the results indicated that the KR group performed as well as the DE group, outperforming the FR group. One of the possible interpretations of this different pattern could be related to the Korean dialect spoken by the participants tested. While the aim of the study was to focus on Seoul Korean, a dialect which has lost any feature that could be associated with lexical stress, some of the participants were either born in other regions of the country or their parents were and still spoke the dialect at home, that is, they were bidialectal (in some cases from Chonnam or North Kyungsang which are claimed to still main a vowel lengthening distinction ([15] [16] [17]). However, excluding these bidialectal participants yielded the same pattern of results, ruling out the possibility that only dialectal differences could explain the results of the KR group.

Other possible interpretations of the results reported should be taken into consideration. While it is possible that KR learners show sensitivity to the greater acoustic saliency of the Stressed fragments, it is also possible that the results obtained are identifying learners' sensitivity to other types of information like lexical information (such as the frequency of the target or competitor word). Although all the efforts were made to match the stimuli (target and competitor) in terms of frequency and length, it was not possible to match the frequency of the words with penultimate and final stress. Words stressed on the penultimate syllable (that is, the target words in the Stressed condition) tend to be more frequent than words with final stress

and represent about 80% of the Spanish lexicon ([22]). Thus, further studies should explore how the frequency of the words and/or of that specific stress pattern or even other variables (such as the amount of lexical competition generated) explain the way in which learners use suprasegmental cues to word recognition in L2 Spanish.

Another possible interpretation of the results, and why all the groups found the Unstressed condition to be particularly challenging, is the fact that this condition does not limit the number of potential lexical candidates that individuals need to consider from within their mental lexicon. For example, when listening to the fragment that contains stress, this cue already limits the number of potential candidates. On the other hand, the Unstressed condition does not constrain the number of potential candidates (the fragment *pelo-* could belong to *pelotón*, but also to *pelonía*, *pelotita*, *pelotillera*, etc.). In line with previous evidence, this could suggest that learners may activate unintended—or phantom—lexical competitors, which may lead to perceptual and, thus, lexical confusion ([29] [30]). Although this point holds validity, it is essential to acknowledge that in real-world scenarios, listeners possess access not solely to bottom-up acoustic phonetic details but also to top-down cues, such as grammatical or contextual information. In this study, the sentences were deliberately designed to be semantically ambiguous, allowing only bottom-up acoustic phonetic information to differentiate between the two forms under scrutiny—one with a vowel and one without. Nevertheless, future research endeavors should investigate whether alternative cues, such as semantic or syntactic cues, interact with bottom-up acoustic phonetic information to limit lexical activation in L2 speech processing.

A final possibility regarding the differences between the FR and KR groups could be related to the prosodic structures of both languages. While both languages are unarguably languages without lexical stress, their prosodic structures are not identical ([13] [14] [28] [31]). While both languages have the AP as the smallest prosodic unit, typically comprising at least a single lexical word and commonly accompanied by preceding (in FR) or following (in KR) clitics ([13] [14]), FR is described as a head/edge language and KR as an edge language whose initial AP tone is delimited by the properties of the first segment ([13] [14]). This discrepancies in the prosodic structures of the AP could point out to the fact that the two languages mark prominence differently which, in turn, could transfer to the perception of L2 lexical stress in different ways.

Finally, and considering that individual variables such as musical aptitude ([5]) , have been found to influence how much individuals struggle with perceiving lexical stress when the property is absent in their L1, future studies should look at how speech perception and word recognition interact with each other, as well as how much different individual variables (L2 proficiency, musical aptitude, or even the knowledge/proficiency in other languages) modulates their use of lexical stress for online word recognition in L2.

5. Conclusions

While future studies still need to explore how variables such as individual differences (e.g., musical aptitude) or the knowledge of other languages with lexical stress influence online word recognition or how the specific properties of the L1 affect the use of suprasegmental cues, this is one of the first studies to investigate and compare how three groups of learners of Spanish (German-speaking, French-speaking, and Korean-speaking learners) make use of stress during word recognition. Results indicated that German and, surprisingly, Korean-speaking listeners are able to make use of cues to stress when the cue is present in the acoustic input, while French listeners could not, which seems to support the existence of a 'stress deafness', at least among this group of learners.

6. References

[1] Soto-Faraco, S., Sebastián-Gallés, N., & Cutler, A. (2001). Segmental and suprasegmental mismatch in lexical access. *Journal of Memory and Language, 45*, 412-432.

[2] Dupoux, E., Pallier, C., Sebastian, N., & Mehler, J. (1997). A destressing "deafness" in French? *Journal of Memory and Language, 36*(3), 406-421.

[3] Dupoux, E., Sebastián-Gallés, N., Navarrete, E., & Peperkamp, S. (2008). Persistent stress 'deafness': The case of French learners of Spanish. *Cognition, 106*(2), 682-706.

[4] Peperkamp, S., & Dupoux, E. (2002). A typological study of stress 'deafness'. *Laboratory phonology, 7*, 203-240.

[5] Schwab, S. & Dellwo, V. (2019). Music and L2 prosody: the role of musical aptitude on the discrimination of stress contrasts. In: Calhoun, S., Escudero, P., Tabain, M., Warren, P. (eds.) *Proceedings of the 19th International Congress of Phonetic Sciences, Melbourne*, Australia 2019 (pp. 1927-1931). Canberra, Australia: Australasian Speech Science and Technology Association Inc.

[6] Tremblay, A. (2008). Is second language lexical access prosodically constrained? Processing of word stress by French Canadian second language learners of English. *Applied Psycholinguistics, 29*(4), 553-584.

[7] 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.

[8] Martínez-García, M. T., Shin, D. J., & Lee, G. (2021). Lexical Effects in the Use of Word-Level Stress for Word Recognition by English and Korean Learners of Spanish. In *Proceedings of the VIII Congreso Internacional de Fonética Experimental CIFE* 2021, Universitat de Girona, Girona, Spain.

[9] Martínez-García, M. T., & Schwab, S. (2023). Relation between Musical Aptitude and L2 Stress Perception in French- and Korean-Speaking Listeners. In *Proceedings of the 20th International Congress of the Phonetic Sciences (ICPhS 2023)*, Prague, Czech Republic.

[10] Lee, G. (2015). *Production and perception of Korean and English word-level prominence by Korean speakers.* Ph.D. Dissertation. University of Kansas.

[11] Lacheret-Dujour, A., & Beaugendre, F. (1999). *La prosodie du français* (French prosody) (CNRS éditions, Paris), pp. 1-354.

[12] Rossi, M. (1979). Le français, langue sans accent (French, language without stress). *Studia Phonetica, 15*, 13-52.

[13] Jun, S.-A., 1993. *The Phonetics and Phonology of Korean Prosody*. Ph.D. dissertation. Ohio State University.

[14] Jun, S.-A., 1998. The accentual phrase in the Korean prosodic hierarchy. *Phonology, 15*(2), 189–226.

[15] Heo, W., 1965. Kwuke Umwoonhak [Korean Phonology]. Cengumsa, Seoul.

[16] Kenstowicz, M., Park, C., 2006. Laryngeal features and tone in Kyungsang Korean: a phonetic study. *Stud. Phon. Phonol. Morphol. 12* (2), 247–264.

[17] Ko, E.-S., 2013. A metrical theory of Korean word prosody. *Ling. Rev. 30*(1), 79–115.

[18] Altenberg, E. P. (2005). The perception of word boundaries in a second language. *Second Language Research, 21*, 325-358.

[19] Weber, M. (2007). DIALANG (review). *The Canadian Modern Language Review, 64*, 233–238

[20] Tagliaferri, B. (2005). *Paradigm*. Perception Research Systems, Inc. Available at www.perceptionresearchsystems.com

[21] 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.

[22] Harris, J. W. (1967). *Spanish phonology* (Doctoral dissertation, Massachusetts Institute of Technology).
[23] R Development Core Team. 2009. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing.
[24] Bates, D., Mächler, M., Bolker, B., & Walker, S. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48.
[25] Baayen, R. H., Davidson, D. J., & Bates, D. M. 2008. Mixed effects modeling with crossed random effects for subjects and items, *J. Memory Lang.*, *59*, 390-412.
[26] Martínez-García, M. T., Van Anne, K., Brown, R., & Tremblay, A. (n.d.). English and Spanish Listeners' Use of "Positive" Stress cues in Spanish Word Recognition.
[27] Kwon, Na-Yeon. (2016). *Learning Stress Cues for Word Segmentation by Korean Learners of English* (Doctoral dissertation, 서울대학교 대학원).
[28] Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of memory and language*, *50*(1), 1-25.
[29] Broersma, M., & Cutler, A. (2011). Competition dynamics of second-language listening. *Quarterly Journal of Experimental Psychology*, *64*(1), 74-95.
[30] Shin, J., & Speer, S. R. (2012). English lexical stress and spoken word recognition in Korean learners of English. In *Speech Prosody 2012*.