The acquisition of Spanish lexical stress by Korean learners

José Ignacio Hualde, Ji Young Kim

University of Illinois at Urbana-Champaign
jihualde@illinois.edu, jkim315@illinois.edu

ABSTRACT

We investigate the acquisition of Spanish lexical stress by native speakers of Korean. Korean is a language without lexical stress, where rising tone cues prominence at the phrase-level. Spanish, has lexical stress. In a first experiment, we examined the production of pitch correlates of lexical stress. Our Korean participants' production was similar to that of native Spanish speakers for sentence-final words, but words in sentence-medial position showed a peak typically aligned near the end of the word, regardless of the position of the stressed syllable. This can be seen as a transfer from native Korean intonational patterns.

In a second study we tested the perception of stress with minimal pairs in different intonational contexts. Overall, learners' accuracy was very low, and with a strong bias for penultimate stress. Performance was slightly better in citation forms.

Keywords: stress, second language acquisition, accent, prosody, Spanish, Korean.

1. INTRODUCTION

We examine the perception and production of Spanish lexical stress by Korean students of Spanish. Whereas Spanish has contrastive lexical stress, Standard Korean does not have any lexically contrastive prosodic properties, as it lacks both lexical stress and lexical tone. Middle Korean was a tone language and lexical tonal contrasts have been preserved in some Korean dialects, although not in Standard Seoul Korean [7]. Unlike in Spanish, stress in Korean is only a phrasal phenomenon. Korean does not have stress contrasts in the lexicon, but syllables realized with rising pitch may be perceived as prominent [6].

Given these important typological differences, the perception and production of lexical stress can be expected to be a challenge for Korean learners of Spanish. A reasonable assumption would be that, at least at an initial stage in their learning, Korean students will interpret Spanish stress in terms of phrasal prosody.

Although most Korean students of Spanish would have studied English before, familiarity with the notion of stress in English would not make the acquisition of Spanish stress unproblematic, since Spanish does not have the systematic reduction of unstressed vowels and other clear segmental correlates of stress that are found in English.

We report on two experiments related to production and perception.

2. PRODUCTION STUDY

2.1. *Research question*

In Spanish, words in final position in a declarative sentence with more than one stressable word typically show either a tonal peak aligned within the stressed syllable or a fall throughout the word [1, 5, 10] . In phrase-initial position, on the other hand, the peak of raising accents is usually displaced to the following syllable. The amount of displacement is affected, in part, by stress pattern, being smallest with oxytones and greatest with proparoxytones [4, 9].

In Korean, non-final accentual phrases, which consist of a lexical item and following particles, are characterized by a pitch rise throughout the phrase, with a peak aligned with its end [6, 8]. Transfer from Korean, should result in peak displacement in prenuclear position, like in native Spanish, but with even greater displacement in proparoxytones, if the right-edge of the word plays a consistent role as a target for peak alignment. At the end of an intonational phrase, on the other hand, Korean shows a final descent in declaratives from a peak on the first or second syllable. Here we focus on words in nonfinal position.

2.2. *Methods*

We have analyzed data from 10 Korean college learners of Spanish (L2: 8 f, 2 m) and 9 native Spanish speakers (NS: 3 f, 6 m).

We selected 20 Spanish words for this production experiment: 10 proparoxytones (e.g., *número* 'number') and 10 paroxytones (e.g., *hermano* 'brother'). The target words were inserted in either final or non nonfinal position in declarative sentences. Here we report on alignment in words in nonfinal (prenuclear) position, e.g. *El número de*

premios es muy limitado. 'The number of awards is very limited.' The sentences were presented to the participants in PowerPoint slides on a computer screen. Participants were asked to read them out loud. The participants' speech was audio-recorded in a sound-attenuated booth, using an AKG C520 head-mounted microphone, which was positioned approximately 2 inches away from each participant's lips, and a Marantz PMD570 solid state recorder.

For all test items, the pitch (in Hz) and time point (in ms.) of f0 peak (H) were semi-automatically coded in *Praat* [2]. The onset and offset and the stressed syllable, and the offset of the target word were manually marked as well.

2.3. *Results and discussion*

A total of 320 tokens of words in nonfinal position were obtained, of which 16 were discarded for difficulties in calculating F0, for a total of 304 tokens entered in the analysis.

Given that H alignment is affected by stressed syllable duration [12, 14, 15], we normalized duration dividing the distance from the F0 peak to the word-edge by the duration of the last syllable.

The effects of participant group (NS/L2) and stress pattern (proparoxytone/paroxytone) on the normalized H-to-Word Offset distance was analyzed using linear mixed effects modeling with subject and item as random factors. Results show that there was a significant main effect of participant group (β = -0.278, SE = 0.0915, t= -3.042), which indicates that in general the distance from the H to the final edge of the word was shorter for the L2s than for the NS. Moreover, there was a significant main effect of stress pattern (β = 0.345, SE = 0.061, t = 5.611) and a significant interaction between group and stress pattern (β = -0.521, SE = 0.076, t = -6.839). This indicates that overall the peak H of the proparoxytones was aligned further away from the word edge than that of the paroxytones and that this difference was smaller for the L2s than the NSs. Pairwise comparisons of group and stress pattern show that while the proparoxytones had longer H-to-Word Offset than the paroxytones for the NS ($p < 0.0001$), no significant difference was found between the L2s' two stress patterns. See Figure 1.

To sum up: the Korean subjects displaced the accentual peak in proparoxytones (e.g. *número*) in prenuclear position significantly further to the right than the native speakers, indicating a tendency to align the high tone with the end of the word both in paroxytones and in proparoxytones. See Figure 2.

Figure 1. Normalized distance from f0 peak (H) to the word edge of paroxytones and proparoxytones of native speakers (NS) and L2 learners (L2). Dotted lines are the onset and offset of the last syllable of the word.

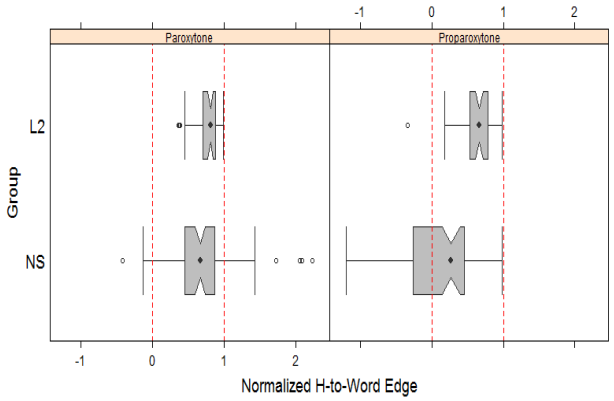
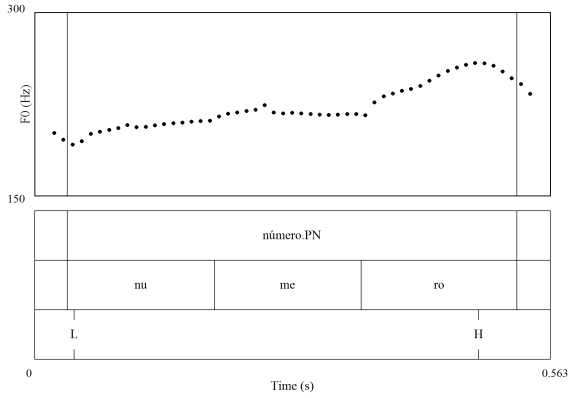


Figure 2. Example of L2 production of *número*, showing peak alignment with the end of the word.



That is, whereas the Spanish native speakers aligned the tonal peak with respect to the stressed syllables, the Korean learners of Spanish appear to have used the end of the word as the landmark. This resulted in much less variation in alignment than for the NSs and no statistical difference betwen paroxytones and proparoxytones in the position of the peak. It is important to note, however, that the L2s did not neutralize the stress contrast, as the beginning of the accentual rise clearly separated paroxytones from paroxytones. Nevertheless the observed difference in the position of the peak with respect to the native pattern raises questions for the perception of native stress contrasts.

3. PERCEPTION STUDY

3.1. *Research question*

In Spanish, stress may fall on any of the last three syllables of the word. Although strict minimal pairs not involving verbs are relatively few, a number of verb forms differ only on the position of the stress, e.g. *canto* 'I sing' vs *cantó* 's/he sang'.

Since subjects of verbs can be left unexpressed in Spanish, the position of the stress in the verbal form may be crucial to interpret meaning. In order to probe the perception of lexical stress we thus focus on verbal minimal pairs. In particular, our question is whether the identification of lexical stress pattern is facilitated or becomes more difficult depending on the intonational context.

3.2. *Methods*

22 Korean college learners of Spanish (L2: 15 , 7 m) and 13 native Spanish speakers (NS: 6, 7 m) participated in the study. According to a Spanish proficiency test (cloze-test), the L2s fell within the low to intermediate proficiency range, with an average of 24.5 out of 50 points (range: 18-43 points). We selected 40 Spanish minimal pairs consisting of two verbal forms for the creation of aural stimuli. The two forms of the pairs are segmentally identical and differ only in the location of the lexical stress, which falls on the penultimate syllable in the present tense form and on the final syllable in the preterit form, e.g. *anudo* 'I tie', vs *anudó* 's/he tied'. 20 of the pairs are disyllabic and the other 20 are trisyllabic. The verbal forms were inserted in four phrasal contexts: C1 Target word in isolation; C2 target word in prenuclear position in statement: (e.g. *anudo la corbata* 'I tie the necktie'); C3 target word in final position in yes/no question (e.g. *¿los anudo?* 'should I tie them?') and C4 target word in prenuclear position in yes/no question (e.g. *¿anudo la corbata?* 'should I tie the necktie?'). These four contexts were chosen because the F0 contour over the target word is expected to be different in each case, based on previous work on Spanish intonation (see, e.g. [3]). In C1 (word in isolation) usually there is a pitch rise and fall confined to the stressed syllable, which should make the identification of the stress pattern of the word easier than in the three other contexts.

To obtain stimuli for our perception experiment, 320 sentences (40 verbs × 2 stress patterns × 4 prosodic contexts) were recorded by one of the authors in a sound-attenuated booth using professional equipment. The contours that were obtained were as expected. For the perception experiment we divided the stimuli into two lists, so that the same listener never heard both members of a stress pair in the same context.

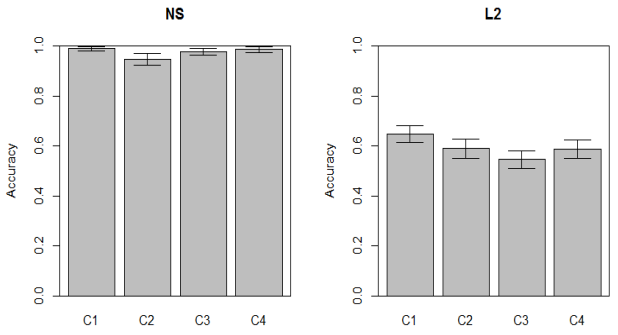
The participants engaged in a forced-choice identification task, in which they listened to the target stimuli through Samsung SHS-250V headphones and determined as fast as they could whether the target verbal form they heard was stressed on the final or on the penultimate syllable by pressing the key of the corresponding subject of the verb (i.e., *él* 'he' for oxytones or *yo* 'I' for paroxytones). The order of options was counter-balanced. The experiment was carried out with the software program PsychoPy2 [11], which also collected participants' response accuracy and reaction times.

Participants' response accuracy was coded as either correct or incorrect. Items with response times that were two standard deviations away from each participant's mean response time were excluded from the analysis. In total, 835 tokens out of 5600 were excluded due to this reason. Response accuracy was statistically analyzed by mixed effects logistic regression using the *glmer*() function in the *lme4* package in R [1]. Following [1], absolute t-values higher than 2 were considered as statistically significant. For post-hoc pairwise comparisons, the *lsmeans*() function in the *lsmeans* package in R [1] was used. Preliminary inspection of the data showed that there were 5 items in the stimuli that the control native speakers identified with less than 60% accuracy. We removed these items for the statistical analysis.

3.3. *Results*

Figure 3 shows accuracy rates of our NSs (left panel) and L2s (right panel) in the perception of stress location in the four prosodic contexts represented in the aural stimuli. As can be seen, the NS group performed well above 90% correct answers in all four contexts (as mentioned before, this is after removing 5 problematic items). For the L2 group instead, accuracy rates were slightly above chance level in all prosodic contexts (range = 54% - 64%).

Figure 3. Accuracy rate in perception task of native Spanish speakers (NS) and Korean L2 learners of Spanish (L2). The four columns in each panel correspond to four phrasal/prosodic contexts. Column height indicates the mean for each group (1.0 = 100% accuracy) and error bars are also shown.

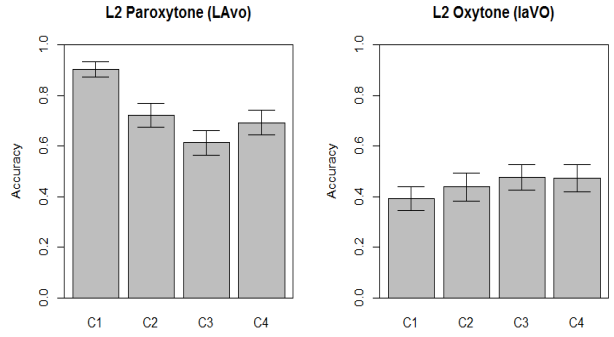


Accuracy scores were submitted to a mixed effects logistic regression with subject and item as random factors. We analyzed the effect of participant group (NS/ L2) and prosodic context, as well as the interaction between these two fixed factors. Stress pattern (paroxytone vs. oxytone) and syllable number (disyllabic vs. trisyllabic) were also included in the model as controlled factors. Through model comparisons, the maximal random effect structure justified by the design was chosen as the best fitting model. All the fixed factors (i.e., group and prosodic context) were centered using contrast-coding.

The regression analysis showed a significant main effect of participant group (β = -4.183, SE = 0.305, z = -12.707, $p < 0.001$), indicating that the NSs (baseline group) performed significantly better than the L2s.

For the L2s, a difference was found between words with final and penultimate stress. Penultimate stress resulted in higher accuracy. See Figure 4.

Figure 4. Accuracy rate in perception task of Korean L2 learners of Spanish (L2) separated by stress pattern of stimuli: paroxytones vs oxytones .



The L2 data were further analyzed in order to determine to what extent stress class significantly affects accuracy in our four sentential contexts. Significant main effects were found for stress pattern (β = 1.83, SE = 0.433, z = 4.23, $p < 0.001$) and prosodic context for S (β = -0.801, SE = 0.353, z = -2.269, $p < 0.05$). That is, the L2s identified the location of stress significantly better in paroxytones than in oxytones. Moreover, their performance was better when the F0 peak aligned with the stress syllable (C1) than when it was not, although the significance of the comparison was limited to one context (C1 vs C2). Significant interactions were also found between stress pattern and C2 (β = -2.517, SE = 0.283, z = -8.899, $p < 0.001$) and stress pattern and C3 (β = -0.596, SE = 0.215, z = -2.775, $p < 0.01$), which indicates that the accuracy difference between paroxytones and oxtyones was significantly larger for context C1 than for C2 and C3. No main effect was found regarding number of syllables.

The L2s' bias toward selecting paroxytones was confirmed by d-prime and C-score analysis. For the L2s, d-prime scores were close to zero (M = 0.498), indicating that they did not have sensitivity in distinguishing the two stress patterns. Their negative C-scores (M = -0.467) confirm that this low sensitivity is due to bias towards responding paroxytone. The NSs, on the other hand, not only had close to perfect sensitivity in distinguishing the two stress patterns, as confirmed in their high d-prime scores (M = 3.971), but also did not show any response bias, with C scores close to zero (M = -0.092).

To sum up, the L2 participants were significantly less accurate than the NSs across all sentential contexts. The stimuli with which they performed best were paroxytones (present tense) in citation context. In general they showed a strong bias to respond "paroxytone" (1st sg present tense form). No difference was found in the accuracy with which they identified two-syllable and three-syllable words.

4. GENERAL DISCUSSION

Our production study shows that our Korean learners of Spanish, like our NSs, anchored a rising pitch accent with the lexically stressed syllable in words in prenuclear position in declarative sentences. An important difference between the L2s and the NSs was nevertheless found in the position of the accentual peak, which for the L2s tended to align with the end of the word both for paroxytones and proparoxytones, showing very little variation in the alignment. Our Korean learners thus implemented the stress contrast using a different strategy than the NS. We hypothesized that if the L2s use nonnative patterns in the production of Spanish stress, this may affect their perception of native speech. This was indeed the case. In an experiment involving a choice between penultimate and final stress, our Korean participants showed a very strong bias towards perceiving stress on the penultimate syllable, to the extent that with oxytonic stimuli they performed at or under chance level in the identification task.

We had hypothesized that L2 participants would do best with words in isolation with declarative intonation. This prediction was confirmed, but only for paroxytones.

7. REFERENCES

[1] Baayen, R. H. 2008. *Analyzing Linguistic Data: A Practical Introduction to Statistics Using R*. Cambridge: UK: Cambridge University Press.

[2] Boersma, P. & Weenink, D. 2014 Praat: doing phonetics by computer [Computer program]. Version 5.4.04, retrieved 28 December 2014 from http://www.praat.org/

[3] Estebas-Vilaplana, E. & Prieto, P. 2010. Castilian Spanish intonation. In: Prieto, P. & Roseano, P. (eds.), *Transcription of Intonation of the Spanish Language*. Lincom Europa: München, 17-48.

[4] Hualde, J.I. 2002 Intonation in Spanish and the other Ibero-Romance languages: Overview and status quaestionis. In Wiltshire, C. & Camps, J. (eds.), *Romance Phonology and Variation*. Amsterdam: Benjamins, 155-184.

[5] Hualde, J. & Prieto, P. to appear. Intonational variation in Spanish: European and American varieties. In: Frota, S. & Prieto, P. (eds.) Intonational variation in Romance. Oxford: Oxford Univ. Press.

[6] Jun, S.-A. 2005. Korean intonational phonology and prosodic transcription. In: Jun, S-A. (ed) *Prosodic Typology: The Phonology of Intonation and Phrasing*. Oxford: Oxford Univ. Press, 201-229.

[7] Kim, K. 2013. *Tone, pitch accent and intonation in Korean: A synchronic and diachronic view*. Doctoral dissertation, Univ. Köln.

[8] Kim, S. 2004. The role of prosodic phrasing in Korean word segmentation. Doctoral dissertation. UCLA.

[9] Llisterri, J., Marin, R., da la Mota, C. & Ríos, A. 1995, Factors affecting F0 displacement in Spanish. *ESCA, Eurospeech'95. 4h Conference on Speech Communication and Technology*, 2061-2064.

[10] O'Rourke, E. 2012. Intonation in Spanish. In Hualde, J.I., Olarrea, A. & and O'Rourke, E. (eds.), *Handbook of Hispanic Linguistics*. Oxford: Blackwell, 173-191.

[11] Pierce, J.W. 2007. PsychoPy - Psychophysics software in Python. *Journal of Neuroscience Methods*, 162(1-2):8-13.

[12] Prieto, P., Estebas-Vilaplana, E., & Vanrell, M. M. 2010. The relevance of prosodic structure in tonal articulation. Edge effects at the prosodic word level in Catalan and Spanish. *Journal of Phonetics* 38: 688-707.

[13] Prieto, P. & Roseano, P. 2010. *Transcription of Intonation of the Spanish Language*. Munich: Lincom Europa.

[14] Prieto, P. & Torreira, F. 2007. The segmental anchoring hypothesis revisited: Syllable structure and speech rate effects on peak timing in Spanish. *Journal of Phonetics* 35: 473–500.

[15] Prieto, P., van Santen, J., & Hirschberg, J. 1995. Tonal alignment patterns in Spanish. *Journal of Phonetics*, 23, 429–451.