SYLLABLE STRUCTURE AFFECTS SECOND-LANGUAGE SPOKEN WORD RECOGNITION AND PRODUCTION

Maria Teresa Martínez-García & Annie Tremblay

University of Kansas
maria.martinezgarcia@ku.edu

ABSTRACT

In this study, we show that second-language (L2) spoken-word recognition is greatly influenced by syllable-structure differences between the native language (L1) and the second language (L2), and that L2 word-recognition accuracy is a reliable predictor of L2 word-production accuracy. Spanish-speaking English learners (experimental group) completed a listening task in which they monitored /(ə)s+Consonant/-initial (henceforth, /(ə)sC/-initial) words in English. Proficiency-matched German-speaking English learners (L2 control group) and native English listeners (L1 control group) completed the same word-monitoring task. The Spanish group also produced the corresponding /(ə)sC/-initial words. The results show a clear effect of L1 on L2 learners' word recognition, with the German group outperforming the Spanish group. For Spanish speakers, a significant positive relationship between word recognition and word production was also observed. These results indicate that L1-L2 syllable-structure differences have pervasive consequences for spoken-word recognition, and L2 word production difficulties may be closely tied to L2 word recognition difficulties.

Keywords: word recognition, syllable structure, phonotactics, consonant clusters, Spanish

1. INTRODUCTION

It is well known that native-language (L1) syllable structure and, more generally, L1 phonotactics influence non-native speech perception. For example, Japanese listeners perceive illusory vowels between consonant sequences that are illegal in the L1 (e.g., [1]). Similarly, Korean second-language (L2) learners of English perceive illusory vowels between word-medial consonant clusters that violate L1 syllable structure constraints (e.g., [2]). These results, which were found in speech perception experiments that used non-words, have been interpreted as evidence that listeners perceptually repair sequences of sounds that are illegal in the L1. By extension, we should also expect L1 syllable structure and, more generally, L1 phonotactics to affect L2 word recognition. Existing research suggests that L2 learners' perceptual difficulties may have important consequences for L2 word recognition by activating unintended (or "phantom") lexical competitors, thus potentially creating lexical confusion (e.g., [3,4]). If L2 learners misperceive syllable structure, they may activate words that native listeners would not necessarily activate.

The present study investigates the degree to which syllable structure affects L2 spoken-word recognition and production. It does so by examining Spanish listeners' recognition and production of /(ə)s+Consonant/-initial (henceforth, /(ə)sC/-initial) words in L2 English. Spanish and English share the segments that form /sC/-initial clusters, but only English allows the two sounds to be combined in syllable- (and thus word-) initial position (e.g., English: *study*; Spanish: *estudiar*). Spanish-speaking L2 learners of English may therefore perceptually repair an /sC/-initial cluster with the insertion of an epenthetic vowel in both spoken word recognition and word production. Unlike Japanese and Korean listeners (e.g., [1,2]), Spanish L2 learners of English have been shown to produce such epenthetic vowels *before* the /s/ rather than between the two consonants, perhaps due to the fact /sC/ sequences in the first two syllables of Spanish words are preceded by the vowel /e/ [5,6,7]. If Spanish listeners indeed perceive such epenthetic vowels in speech, they should activate unintended competitor words (e.g., *especially*) rather than the intended target words (e.g., *specially*), making L2 word recognition more difficult (e.g., [3,4]). This perceptual difficulty may also make them less likely to produce /sC/-initial words accurately.

Using a word monitoring task (Experiment 1), we compared the performance of Spanish-speaking L2 learners of English to that of proficiency-matched German-speaking L2 learners of English. Although German does not allow /sC/-initial clusters, it does allow /ʃC/-initial consonant clusters (e.g., *Straße* [ˈʃtʁaːsə] 'street'). Thus, at the level of syllable structure, German and English allow a similar combination of sounds in onset position (i.e., sibilant+Consonant) that Spanish does not allow. If German listeners were to have difficulty recognizing /sC/-initial words in English due to phonotactic differences in the precise sound combinations allowed in the two languages, they would be unlikely

to experience interference from /əsC/-initial words, which were the lexical competitor words used in the present study. German listeners will thus serve as the L2 control group to which Spanish listeners will be compared. This comparison will help determine the degree to which L1-L2 differences in syllable structure affects L2 spoken word recognition. We predict that German-speaking L2 learners of English will outperform proficiency-matched Spanish-speaking L2 learners of English in the recognition of /(ə)sC/-initial words.

A second issue that this study considers is the degree to which L2 word production difficulties are related to L2 word recognition difficulties. Spanish-speaking L2 learners of English are known to produce epenthetic vowels before /sC/-initial clusters in their oral productions (e.g., [5,6,7]). To test the relationship between L2 word production and L2 word recognition, we asked the same Spanish-speaking L2 learners of English to produce the sentences they heard in the word recognition task (Experiment 2), and we examined whether L2 word recognition accuracy was a reliable predictor of L2 word production accuracy.

2. EXP. 1: WORD MONITORING TASK

2.1. Participants

Thirty-two native speakers of English (mean age=23 years), 32 Spanish-speaking L2 learners of English (mean age=24 years), and 32 German-speaking L2 learners of English (mean age=23 years) participated in this study. The L2 groups were matched in terms of age of acquisition ($t<|1|$), months of residence ($t(62)=-1.77$, $p>.1$), years of instructions ($t<|1|$), and proficiency as established by a cloze (i.e., fill-in-the-blank) test ($t(62)= -1.12$, $p>.1$) [8]. All L2 learners had first been exposed to English after the age of 9.

2.2. Materials

Forty-eight /(ə)sC/-initial experimental pairs (36 minimal, 12 near-minimal) (e.g., *specially* vs. *especially*) were selected for the task. To obtain a balanced design, participants saw either an /sC/-initial word or an /əsC/-initial word on the computer screen, and they heard a semantically neutral sentence that contained either the /sC/-initial word or the /əsC/-initial word, resulting in the 2 × 2 design illustrated in Table 1. The sentences were created so that the critical word would appear in different positions in the sentence (initial, medial, or final) and never after a word-final /s/ or /ə/. The experiment also included 96 filler sentences where participants monitored other types of words. All sentences were checked by two (other) native English speakers to ensure they were plausible and neutral between the two words of any given pair. The sentences were recorded by a female native speaker of American English. Items were presented in a Latin Square design such that participants did not see/hear any item in more than one condition. Test items were randomized across participants.

Table 1: Example of stimuli used in the task.

Correct Response	Written Target	Auditory Stimulus
'yes' (match)	*SPECIALLY*	*I prepared that specially for you.*
	ESPECIALLY	*I prepared that especially for you.*
'no' (mismatch)	*SPECIALLY*	*I prepared that especially for you.*
	ESPECIALLY	*I prepared that specially for you.*

2.3. Procedures

The experiment was administered in Paradigm (Perception Research Systems, Inc. [9]). In each trial, participants saw the target word appear in the middle of the screen in capital letters; the word was present on the screen for 1,000 ms and disappeared as the auditory stimulus started playing. Participants were asked to decide if the sentence contained the word they saw on the screen by pressing 'yes' (i.e., the left button of the mouse) as soon as they recognized the target word in the sentence or by pressing 'no' (i.e., the right button of the mouse) after the end of the auditory sentence. The next trial began immediately after participants entered their response. A practice session of six stimuli with feedback preceded the main experiment.

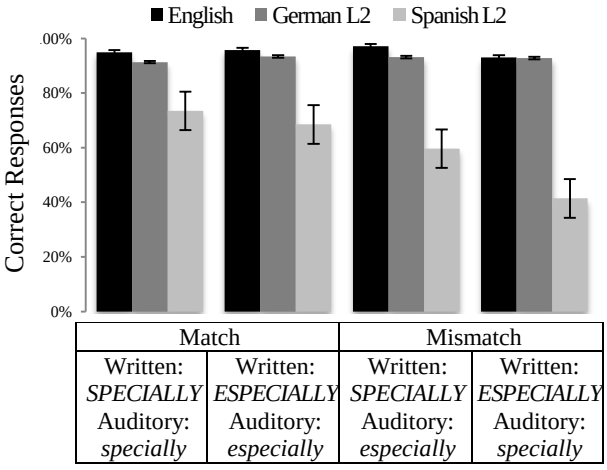
2.4. Data Analysis

Participants' accuracy was analyzed with logit mixed-effects models (cf. [10]). A first model on all the groups' accuracy included the following fixed variables: L1 (English, Spanish German; baseline=English), match between the written word and the word in the auditory stimulus (match vs. mismatch; baseline= match), initial vowel in the written word (presence, absence; contrast coded presence= –0.5). Following significant interactions, two follow-up models were conducted on only the Spanish group's accuracy separately for the match and mismatch conditions. Since proficiency did not improve these two models, we report the models without proficiency. All models included participant and test item as crossed random variables.

2.5. Results

Figure 1 presents the percentage of correct responses of all three L1 groups in each of the four conditions.

Figure 1: Percentage of correct responses in the word monitoring task



Match		Mismatch	
Written: *SPECIALLY* Auditory: *specially*	Written: *ESPECIALLY* Auditory: *especially*	Written: *SPECIALLY* Auditory: *especially*	Written: *ESPECIALLY* Auditory: *specially*

A first logit mixed-effects model revealed the following significant effects: L1 for the Spanish group (z(4608)=–14.01, p<.001), who was less accurate than the English group; L1 × initial vowel for the Spanish group (z(4608)=2.29, p<.05), who, unlike the English group, performed less accurately when the written word began with a vowel than when it began with "S"; and an L1 × match × initial vowel interaction for the Spanish group (z(4608)=–2.68, p<.01), who, unlike the English group, showed a greater effect of initial vowel in the mismatch condition than in the match condition. No other effects were significant. These results indicate that only Spanish L2 learners of English had difficulty recognizing /(ə)sC/-initial words.

Given these significant interactions, two follow-models on the Spanish group's accuracy revealed significant effects of vowel for both the match condition (z(768)=2.41, p<.02) and the mismatch condition (z(768)=4.12, p<.001), with Spanish listeners obtaining poorer accuracy when the written word began with a vowel than when it began with "S". The three-way interaction obtained in the first model thus comes from the greater effect of mismatch when the written word began with a vowel than when it began with "S". This indicates that when a competitor with an initial vowel is activated, Spanish L2 learners of English have more difficulty determining whether the word in the auditory stimulus contained a vowel.

2.6. Discussion

The results of the word-monitoring task showed that both native English speakers and German L2 learners of English could correctly detect the target word in the auditory stimuli, independently of the presence or absence of a vowel in the written word or in the auditory stimulus. By contrast, Spanish L2 learners of English were much less accurate in detecting the target word, particularly when there was a mismatch between the word they saw and the word they heard in the sentence and when the word they saw began with a vowel. These results suggest that L1-L2 differences in syllable structure have pervasive consequences for L2 word recognition.

We now turn to the word production task and examine whether L2 word recognition accuracy is a reliable predictor of L2 word production accuracy. Only the Spanish group completed the word production task.

3. EXP. 2: PRODUCTION TASK

3.1. Participants

The same 32 Spanish-speaking L2 learners of English took part in Experiment 2. They completed the word monitoring task (Exp. 1) before the word production task (Exp. 2). This ensured that any possible awareness of the word types used in the production task would not affect their word recognition results.

3.2. Materials and Procedures

The materials were the same as those described in Experiment 1, including filler items. Participants were audiorecorded while they read aloud the same sentences they had heard in the word monitoring task. Sentences were presented in a similar Latin Square design and were randomized across participants.

2.3. Data Analysis

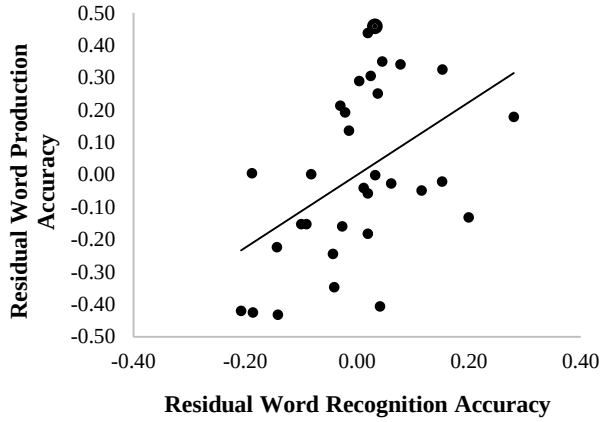
The L2 learners' productions were extracted from the sentences. All /sC/-initial words were analyzed acoustically. Based on the waveform and spectrogram, word productions were coded as accurate (1) if they did not contain any trace of a vowel; otherwise, they were coded as inaccurate (0). The L2 learners' production accuracy was predicted by their word recognition accuracy using a fit linear model that used generalized least squares (cf. [10]). Since proficiency (i.e., cloze test scores) was a significant predictor of both word recognition accuracy (r^2=0.18) and word production accuracy (r^2=0.13), residual word recognition and word production scores (i.e., scores that are not already

explained by proficiency) were calculated and used in the fit linear model. Participant and item were included as random variables.

2.4. Results

Figure 2 illustrates the relationship between L2 learners' residual word recognition accuracy and their residual production accuracy. The results revealed a significant effect of word recognition accuracy ($t(768)=3.44$, $p<.01$), which indicates that word recognition accuracy is a reliable predictor of word production accuracy.

Figure 2: Residual accuracy in word production as predicted by residual accuracy in word recognition.



2.5. Discussion

The results of the word production task showed that Spanish listeners' (residual) word production accuracy can be predicted from their (residual) word recognition accuracy. This suggests that word production difficulties may be closely tied to word recognition difficulties, even after L2 proficiency has been accounted for.

3. GENERAL DISCUSSION

The present study examined the degree to which L1-L2 syllable structure differences affect L2 word recognition and L2 word production.

We first compared Spanish-speaking L2 learners of English and proficiency-matched German-speaking L2 learners of English in their ability to recognize /(ə)sC/-initial words in speech. Spanish and English share the segments that form /sC/-initial clusters, but only English allows the two sounds to be combined in syllable- (and thus word-) initial position. Although German does not allow /sC/-initial clusters, it allows /ʃC/-initial clusters. Thus, at the level of syllable structure, German and English allow a similar sound combination in onset position (sibilant+Consonant) that Spanish does not allow.

The results of Experiment 1 indicated that the Spanish L2 learners of English were much less accurate than both the native English listeners and the German L2 learners of English in recognizing /(ə)sC/-initial words in speech, particularly when there was a mismatch between the word they saw and the word they heard and when the word they saw began with a vowel. These findings, together with those of previous studies (e.g., [3,4]), suggest that non-native listeners' perceptual difficulties have pervasive consequences for L2 spoken word recognition, resulting in the activation of unintended (or "phantom") lexical competitors and important creating lexical confusion for L2 learners.

One limitation of Experiment 1 is that due to the limited availability of /(ə)sC/-initial word pairs in English, we could not control for the lexical frequency of each word in the pair, nor could we use only minimal pairs. The effect of initial vowel observed in the experiment may thus be partly due to lexical frequency (/sC/-initial words were more frequent than /əsC/-initial words). Despite this limitation, the accuracy difference obtained between the Spanish group and the English and German groups is striking, suggesting that Spanish L2 learners of English have difficulty in recognizing words that violate L1 syllable structure constraints.

This study also examined whether Spanish listeners' L2 word recognition accuracy was a reliable predictor of their L2 word production accuracy. /sC/-initial words are known to be difficult to produce for Spanish L2 learners of English (e.g., [5,6,7]). The results of Experiment 2 revealed a significant relationship between L2 learners' word recognition accuracy and their word production accuracy, suggesting a strong tie between L2 word recognition and L2 word production.

One question that arises from these results is whether L2 learners' perceptual difficulties, evidenced in their word recognition results, are in part responsible for their production difficulties. If spoken words are not perceived accurately, a logical consequence should be that they would also not be produced accurately. At the same time, accurate perception should not necessarily entail accurate production, in that articulatory problems may well underlie some of L2 learners' word production difficulties. Our results are most straightforwardly explained by an account that assumes that both L2 word recognition and L2 word production depend in part on a common set of representations (see [11]). Further research should examine the precise nature of these representations.

4. REFERENCES

[1] Dupoux, E., Pallier, C., Kakehi, K., Mehler, J. (1999). Epenthetic vowels in Japanese: a perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance, 25*, 1568–1578.

[2] Kabak, B., Idsardi, W. J. (2007). Perceptual distortions in the adaptation of English consonant clusters: Syllable structure or consonantal contact constraints? *Language and Speech, 50*, 23–52.

[3] Broersma, M., Cutler, A. (2011). Competition dynamics of second-language listening. *Quarterly Journal of Experimental Psychology, 64*, 74–95.

[4] Weber, A., Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language, 50*, 1–25.

[5] Abrahamsson, N. (1999). Vowel epenthesis of /sC (C)/onsets in Spanish/Swedish interphonology: longitudinal case study. *Language learning, 49*, 473–508.

[6] Carlisle, R. (1997). The modification of onsets in a markedness relationship: Testing the interlanguage structural conformity hypothesis. *Language Learning, 47*, 327–361.

[7] Carlisle, R. S. (1998). The acquisition of onsets in a markedness relationship. A longitudinal study. *Studies in Second Language Acquisition, 20*, 245–260.

[8] Brown, J. D. (1980). Relative merits of four methods for scoring cloze tests. *Modern Language Journal, 64*, 311–317.

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

[10] Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge University Press.

[11] Shaw, J., Davidson, L. (2011). Perceptual similarity in input-output mappings: a computational/ experimental study of non-native speech production. *Lingua 12*, 1344–1358.