

# Orthographic Effects in Word Recognition among Spanish-Speaking Learners of English

María Teresa MARTÍNEZ GARCÍA

<https://orcid.org/0000-0003-3187-0853>; [maria.martinezgarcia@utah.edu](mailto:maria.martinezgarcia@utah.edu)  
University of Utah Asia Campus (SOUTH KOREA)



© author's

Suggested citation: MARTÍNEZ GARCÍA, M<sup>a</sup> T. (2023), Orthographic effects in word recognition among Spanish-speaking learners of English, *Langue(s) & Parole*, 8, 127-148 <https://doi.org/10.5565/rev/languesparole.131>

## Abstract

While speech perception has been found to influence word recognition, what specific aspects of the L1-L2 mapping play the most important role is not well understood. This study explores whether, and if so how, visual-orthographic information influences the mapping of phonemic information in an L2 in both perception and word-recognition. Spanish-speaking English learners completed an AXB and a word monitoring task in English manipulating the presence of the labiodental /v/ and the bilabial /b/ phonemes, which are allophonic in Spanish, but phonemic in English (e.g., *best* vs. *vest*). The results show a clear effect of L1 on L2 learners' perception and word recognition, partly modulated by the orthographic information seen. These results support models that predict the mapping of L1-L2 phonemes, emphasizing how orthographic information may be an important variable to take into consideration.

**Keywords:** bilabial; labiodental; L2 mapping; orthographic information; word recognition; speech perception

## Résumé

Bien que la perception de la parole influence la reconnaissance des mots, on ne comprend pas bien quels aspects spécifiques de la cartographie L1-L2 jouent le rôle le plus important. Cette étude explore si les informations d'orthographe visuelle influencent la cartographie des informations phonémiques dans une L2 et, le cas échéant, quelle en est l'incidence à la fois sur la perception et sur la reconnaissance des mots. Un groupe d'apprenants hispanophones d'anglais ont effectué une tâche AXB et une tâche de monitoring de mots anglais dans lesquels la place des phonèmes labiodental /v/ et bilabial /b/, qui sont allophoniques en espagnol mais phonémiques en anglais (par exemple, *best* vs. *vest*), avait été modifiée. Les résultats montrent un effet clair de la L1 des apprenants

sur la perception et la reconnaissance des mots de la L2, modulé en partie par les informations orthographiques vues. Ces résultats soutiennent les modèles qui prédisent la cartographie des phonèmes L1-L2, en soulignant à quel point les informations orthographiques peuvent être une variable importante à prendre en compte.

**Mots clés :** bilabial ; labiodental ; cartographie L2 ; informations orthographiques ; reconnaissance de mots ; perception de la parole

### Resumen

Si bien se ha descubierto que la percepción del habla influye en el reconocimiento de palabras, no se comprende bien qué aspectos específicos del mapeo L1-L2 desempeñan el papel más importante. Este estudio explora si la información visual-ortográfica influye en el mapeo de la información fonémica en una L2, y de ser así, cómo, tanto en la percepción como en el reconocimiento de palabras. Un grupo de hispanohablantes estudiantes de inglés completó una tarea de AXB y una de monitoreo de palabras en inglés en las que se manipuló la presencia de los fonemas labiodental /v/ y bilabial /b/, que son alofónicos en español, pero fonémicos en inglés (p. ej., *best* vs. *vest*). Los resultados muestran un claro efecto de la L1 de los estudiantes en la percepción y el reconocimiento de palabras de L2, modulado en parte por la información ortográfica vista. Estos resultados respaldan modelos que predicen el mapeo de fonemas L1-L2, enfatizando cómo la información ortográfica puede ser una variable importante que debe tenerse en cuenta.

**Palabras clave:** bilabial; labiodental; mapeo L2; información ortográfica; reconocimiento de palabras; percepción del habla

### Resum

Tot i que s'ha descobert que la percepció de la parla influeix en el reconeixement de paraules, no es comprèn bé quins aspectes específics del mapatge L1-L2 fan el paper més important. Aquest estudi explora si la informació visual-ortogràfica influeix en el mapeig de la informació fonèmica en una L2 i, si és així, com, tant en la percepció com en el reconeixement de paraules. Un grup d'estudiants hispanoparlants d'anglès va completar una tasca d'AXB i una de monitorització de paraules en anglès en les quals es va manipular la presència de fonemes labiodental /v/ i bilabial /b/, que són al·lofònics en espanyol, però fonèmics en anglès (p. ex., *best* vs. *vest*). Els resultats mostren un efecte clar de la L1 dels estudiants en la percepció i el reconeixement de paraules de L2, modulats en part per la informació ortogràfica vista. Aquests resultats donen suport a models que prediuen el mapeig de fonemes L1-L2, emfatitzant com la informació ortogràfica pot ser una variable important que cal tenir en compte.

**Paraules clau:** bilabial; labiodental; mapatge L2; informació ortogràfica; reconeixement de paraules; percepció de la parla

## 1. Introduction

An increasingly extensive body of research has demonstrated the influence of the orthographic forms (spellings) of a second language (L2) on speech perception, production, and sound categorization among L2 speakers. Numerous studies, such as those recently conducted by Bassetti and colleagues, have shown that L2 speakers can even produce sound contrasts that do not exist in the target language, but that this effect is modulated by visual-orthographic saliency (e.g., Bassetti, 2017; Bassetti *et al.*, 2015; 2018; 2020). This study focuses on how speech perception and categorization problems might influence word recognition among Spanish learners of English as an L2, and how such an effect may be modulated by orthographic information.

The orthography of both English and Spanish has the graphemes <b> and <v>, but only in English they are distinguished phonetically. While in Spanish both graphemes are pronounced as [b] or [β], depending on the phonetic context and regional variation, English <v> is realized as a labiodental and <b> as a bilabial (Hualde, 2014; Jogman, Wayland, Wong, 1998). In standard Spanish, the voiced labiodental fricative does not exist. However, Hualde (2014) reported [v] in Spanish as a result of coarticulation, such that a word as *Afgano* (Afghan) would be pronounced as [av'ɣano]. This difference in the phonemic repertoire of English and Spanish (yet similar in their orthographic representations) will be crucial for the current study.

The presence of the voiced labiodental fricative sound [v] in the speech of Spanish-English bilinguals, particularly Spanish adult learners of English as an L2, has been discussed in the literature on Spanish-English bilingual speakers. One viewpoint suggests that the presence of the phoneme /v/ in English influences the Spanish sound system, causing [v] to appear as a variant of the phoneme /b/. In other words, the knowledge of English and the speaker's proficiency in English affect how they realize the Spanish sound contrast (e.g., Takawaki, 2012; Trovato, 2017). This influence is seen in Spanish-English bilingual communities, mainly in areas like New Mexico and California, where there is significant language contact between English and Spanish, but the effect varies depending on proficiency in English (e.g., Takawaki, 2012; Tim, 1976; Torres Cacoullós,

Ferreria, 2000). Stevens (2000) also noted [v] in the speech of Spanish L2 instructors, which was influenced by English orthography. Overall, it seems that the presence of the voiced labiodental fricative /v/ in Spanish-English bilingual speech can be attributed to the influence of English on the Spanish sound system, especially in contexts with significant language contact between the two languages and with increasing proficiency in the L2.

Not only direct language contact and proficiency, but also the impact of orthographic forms (spellings) on speech perception in both native and L2 learners has been reported in the literature. Visual-orthographic input has been shown to both facilitate and hinder speech perception. In some cases, providing alphabetic orthographic input alongside auditory input aids speech perception (e.g., Erdener Burnham, 2005; Escudero *et al.*, 2008), but it can also lead to misperceptions if there are discrepancies between the spoken and written forms (e.g., Hayes-Harb *et al.*, 2010; Escudero, Wanrooij, 2010, Experiment 2; Mathieu, 2016; Rafat, Stevenson, 2019). In most instances, the effect reported has been shown in different ways, causing additions, omissions, and substitutions of sounds in speech perception. Additionally, the findings in L2 speech production studies support the evidence from perception studies, as orthographic forms can result in sound modifications in L2 speech (for a review, see Bassetti *et al.*, 2015). Surprisingly, there is a lack of research on whether orthography affects word recognition, despite its influence on speech perception, production, and metalinguistic awareness in L2 learners.

This study, then, focuses on L2 learners' acquisition of non-native phonemic contrasts and investigates how potential perception issues may affect word recognition, and how this pattern could be linked to the visual influence of orthography.

## **2. Experiment 1: AXB Task**

### **2.1 Participants**

32 native speakers of English (16 females; mean age=23 years) and 32 Spanish-speaking learners (12 females; mean age=24 years) participated voluntarily in this study. The native speakers were students at a midwestern university in the USA. The L2 learners were tested at the same institution

(20) and in Spain (12). On average, learners had acquired their L2 at the age of 12 (std=2.3 years), had learned the language for 12 years (std=3.6 years), resided in an English-speaking country for 15 months (std=36.8 months) and had a proficiency, as established by a cloze test (Brown, 1980), of 25.94 out of 50 possible points (std=12.04 points). None of the participants reported having any visual nor hearing impairment.

## 2.2 Materials

Participants completed an AXB discrimination task. Eighteen nonce word minimal pairs were created contrasting the /b/ and /v/ phonemes in English, to avoid participants using lexical information. Each stimulus pair either contained a voiced bilabial stop (e.g., /bemɪʃ/) or a voiced labiodental fricative (e.g., /vɛmɪʃ/), with the contrast appearing in different positions within the word.

The study also included seventy-two fillers. These nonce word fillers were divided in two conditions: The first filler condition included stimulus pairs that contained either a schwa (e.g., /əslɛn/) or did not contain a schwa (e.g., /slɛn/) before the consonant cluster; the second filler condition included stimuli that differed in the presence of one phoneme (e.g., /snu/ vs. /snu\_/). To make sure participants were paying close attention to the complete word, the sound contrasts of these two filler conditions appeared in different positions in the word (initial, middle, or final).

The words used were checked by a native English speaker to ensure they followed English phonotactic rules and were not real English words. To avoid participants relying on the physical properties of the sounds, the stimuli were recorded by three different speakers. The order of presentation was consistent, with Speaker 1 (Midwest dialect) producing A, Speaker 2 (East Coast dialect) producing X, and Speaker 3 (Midwest dialect) producing B. The experiment was designed in a Latin square pattern to prevent participants from hearing both A and B as X.

## 2.3 Procedures

The stimuli were presented using Paradigm by Perception Research Systems, Inc. (Tagliaferri, 2005). Participants were comfortably seated in a quiet room, and they were instructed in their native language to listen carefully to a series of three nonce words and to choose whether the second word (X) was more similar to the first or to the third word (A or

B). Participants were seated comfortably at about 20-30 cms. from the screen and materials were presented in Arial (20 pts.) font. The inter-stimulus interval was 1,000 ms and participants made their decision by pressing one of the two buttons of the mouse. The next trial started immediately after the participants entered their response. A practice session of six stimuli with feedback preceded the main session of the experiment (which did not have feedback). All trials were randomized across participants.

## 2.4 Data Analysis

The participants' accuracy was analyzed with a logistic regression model (cf. Baayen, 2008), using the *glm* package (Hothorn, Everitt, 2014) in R (R Development Core Team, 2009). L1 was considered as a categorical predictor with two levels (English vs. Spanish), with the English group representing the baseline. Phoneme type was considered as a categorical predictor with two levels (Bilabial vs. Labiodental), with the bilabial representing the baseline, as both languages have the bilabial phoneme in their L1s. The effect of the predictors was assessed using log-likelihood tests comparing models with and without them. Two sets of models were run—one on all the accuracy rates with L1 and phoneme type as predictors (and its interaction), and one on the L2 learners' accuracy rates with phoneme type, proficiency, and its interaction. The effect of L2 proficiency was assessed by comparing models that included proficiency with models that did not include it; in each case, the model with the best fit was kept. Since proficiency did not improve the model, only the analysis of all the accuracy data with L1 and phoneme type (but not its interaction) is reported. Participants and item were included as random variables.

## 2.5 Results

Figure 1 presents the mean accuracy results for the two groups in the two phoneme conditions, and Table 1 presents the results of the logit mixed-effects model for all participants' accuracy.

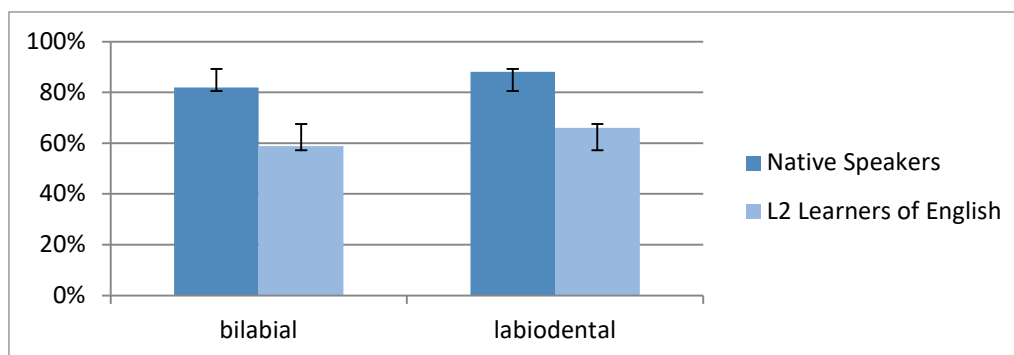


Figure 1: Mean accuracy (standard errors) of the two groups in the AXB task

| Variable                  | Estimate (SE) | <i>z</i> | <i>p</i> |
|---------------------------|---------------|----------|----------|
| (Intercept)               | 1.63 (.18)    | 9.16     | <.001    |
| L1: Spanish L2 Learners   | -1.29 (.19)   | -6.82    | <.001    |
| Phoneme Type: Labiodental | 0.4 (.16)     | 2.55     | <.05     |

Note: *df* = 972;  $\alpha$  = .05

Table 1: Logit regression model on all participants' accuracy results

The model summarized in Table 2 revealed that Spanish L2 learners of English were statistically less accurate than native speakers in discriminating stimuli with [b] and [v] sounds, and that the labiodental sound was statistically more easily identifiable as compared with the bilabial sound by both groups similarly.

## 2.6 Discussion

In this first experiment, an AXB task was used to examine whether Spanish L2 learners would have any difficulty in the discrimination of the bilabial and labiodental contrast in English. Results showed that native speakers were more accurate than the L2 learners and that the labiodental sound was perceived more accurately than the bilabial sound. However, the lack of interaction between the group condition and the phoneme type condition indicates that, while native speakers were more accurate, L2 learners showed the same pattern of results independently whether they were listening to a bilabial (existing in their L1 as a phoneme) or a labiodental (not existing in their L1 as a phoneme, but as an allophone of /f/) or a labiodental sound (not existing in their L1).

The results of this discrimination task for the two groups were in accordance with the phonemic inventory of their L1s. The Spanish

speakers' perception of the stimuli was shaped to conform to the allophonic nature of the [b] vs. [v] distinction in their L1. Considering that the two sounds are allophonic (in very limited contexts, as seen in the introduction) in Spanish, these L2 learners of English have some issues in order to discriminate between them.

These findings add support to the body of evidence that suggests that contrasts that are allophonic in the L1 are, indeed, difficult to discriminate in a contrast which is phonemic in the L2 (e.g., Abramson, Lisker, 1970; Goto, 1971; Lisker, Abramson, 1967; Polka, 1992; Polka, Werker, 1994; Strange *et al.*, 2001; Veleva, 1985; Werker Lalonde, 1988; Werker, Logan, 1985; Werker, Tees, 1984). Yet, it is yet to be established how these perception errors influence L2 word recognition. Existing literature indicates that challenges in accurately perceiving L2 sounds can lead to heightened lexical competition, resulting in less efficient word recognition (e.g., Broersma, Cutler, 2011; Escudero, 2007; Martínez-García, 2021; Weber, Cutler, 2004). Experiment 2 was thus created to examine how the misperception of this phonemic contrast in English could impact L2 word recognition and how orthographic information would play a role in the perception of this sound contrast.

### **3. Experiment 2: Word Monitoring Task**

#### **3.1 Participants**

In Experiment 2, the same participants who were involved in Experiment 1 took part. They performed the word monitoring task before the AXB task to prevent any potential influence on their word recognition from what they might have noticed in the AXB task.

#### **3.2 Materials**

For this experiment, all groups completed a word monitoring task, which requires participants to keep track of a pre-designated target word in the acoustic input (see Kilborn, Moss, 1996). Participants were asked to monitor a target word either containing or not the target word (i.e., *best* vs. *vest*) in semantically ambiguous sentences that could (or could not) contain that given sound (i.e., match vs. mismatch the target word). Table 2 illustrates these four conditions with an example.



| Condition                     | Written Target | Auditory Stimulus           |
|-------------------------------|----------------|-----------------------------|
| Match<br>Stimulus with /b/    | BEST           | I gave her my <u>best</u> . |
| Match<br>Stimulus with /v/    | VEST           | I gave her my <u>vest</u> . |
| Mismatch<br>Stimulus with /b/ | BEST           | I gave her my <u>vest</u> . |
| Mismatch<br>Stimulus with /v/ | VEST           | I gave her my <u>best</u> . |

Table 2: Example of stimuli used in the word monitoring task

The 48 target items included had either the phoneme /b/ or /v/ in different positions within the word (e.g., "bail" vs. "veil"). To prevent participants from relying on lexical cues, semantically ambiguous sentences were created and checked for plausibility and ambiguity by two native English speakers. Additionally, the location of the target word within the sentence was manipulated to prevent participants from forming expectations.

The experiment also included ninety-six filler items, which shared characteristics with the fillers used in Experiment 1. Half of these fillers consisted of minimal (36) and near-minimal (12) pairs in which the presence of a vowel at the beginning of /s/-initial clusters was manipulated (e.g., "state" vs. "estate"). The other half of the fillers differed in the number of phonemes they contained or in one of their phonemes (e.g., "stop" vs. "top" or "snow" vs. "know"). These fillers had sound contrasts appearing in different positions in the word (word-initial, word-medial, or word-final). All the sentences, both experimental and fillers, were recorded by a female native speaker of American English with a Midwestern accent (Speaker 3 of Experiment 1). A randomized Latin square design was used to present the stimuli.

### 3.3 Procedure

The presentation of the stimuli was done using Paradigm. In each trial, participants saw the target word in the middle of the screen in capital letters and for 1,000 ms. (e.g., BEST). The word disappeared at the same time as the audio started playing. Then, they listened to a sentence that might or might not have included the word they just saw (e.g., *I gave her my*

*best* or *I gave her my vest*). Participants' task was to decide whether the sentence contained the word they saw on the screen by pressing the button "SÍ", or otherwise press "NO" in a mouse. Participants could make the decision as soon as they identified the word in the sentence (before finishing the audio) or, otherwise, wait to the end of the sentence to make sure the word did not appear otherwise. As soon as they entered their response, the next trial started.

The experiment started with six practice trials, which included feedback. During the main session of the experiment, participants did not receive any feedback.

### 3.4 Data Analysis

The participants' accuracy in this task was analyzed using a logistic regression model as in Experiment 1. Three variables were considered as categorical predictors. First, L1 with two levels (English vs. Spanish), with English serving as the baseline. Second, the effects of the phoneme type (e.g., /b/ vs. /v/), with the bilabial sound as the baseline. Finally, the match between the word to be monitored and the word in the auditory stimulus (match vs. mismatch), with "Match" serving as the baseline.

In this study, two sets of models were run. The first set used L1 as a predictor and considered all participants' accuracy rates. The second set focused on L2 learners and their accuracy rates, along with their proficiency. However, the results regarding L2 proficiency are not reported because the model without proficiency provided the best fit. Both sets of models included participant and item as random variables in their analysis.

### 3.5 Results

Figure 2 presents the mean accuracy of the two groups in the word monitoring task, and Table 3 presents the results of the logistic regression model for all participants' accuracy.

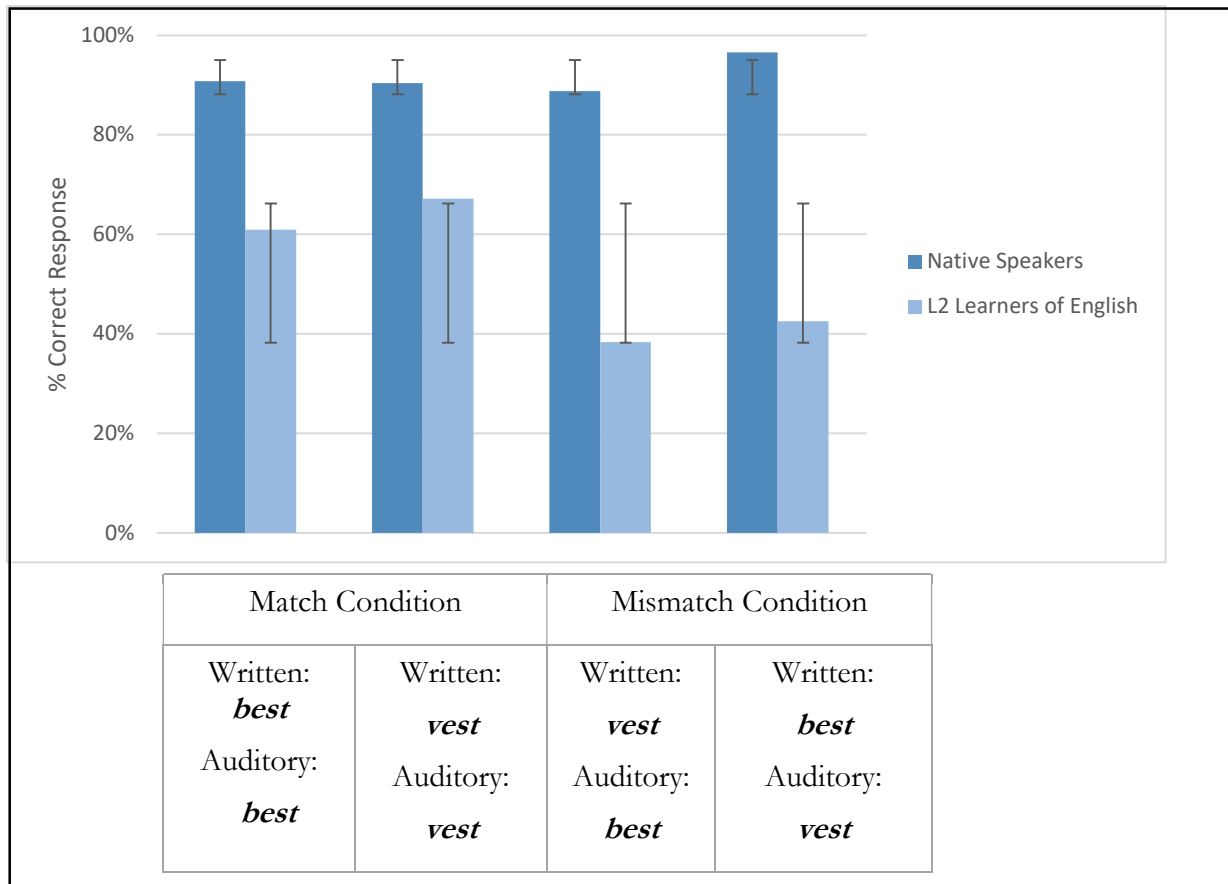


Figure 2: Mean accuracy (standard deviation) of the three L1 groups in the word monitoring task

| Variable                          | Estimate (SE) | <i>z</i> | <i>p</i> |
|-----------------------------------|---------------|----------|----------|
| (Intercept)                       | 2.41 (.22)    | 11.01    | <.001    |
| L1: Spanish L2 Learners           | -1.92 (.26)   | -7.4     | <.001    |
| Auditory Type: Mismatch           | .23 (.27)     | -.85     | >.1      |
| Phoneme Type: Labiodental         | -.05 (.27)    | -0.17    | >.1      |
| L1 x Auditory Type                | -0.77 (.31)   | -2.46    | <.05     |
| L1 x Phoneme Type                 | 0.34 (.31)    | 1.08     | >.1      |
| Auditory Type x Phoneme Type      | 1.34 (.46)    | 2.93     | <.01     |
| L1 x Auditory Type x Phoneme Type | -1.45 (.51)   | -2.84    | <.01     |

Note: *df* = 2592;  $\alpha$  = .05

Table 3: Logit regression model on all participants' accuracy results

In the results of the model, the following effects can be observed: A main effect of L1, indicating that learners showed a different pattern compared with the native speakers; an interaction between auditory type and L1 for Spanish L2 learners, which shows that, unlike the native

speakers, these learners showed different accuracies when the words contained /b/ vs. /v/; an interaction between auditory type and phoneme type, which indicates that bilabials and labiodental sounds showed different accuracies depending on whether they were presented in the match or mismatch condition; and a three-way interaction between phoneme type, auditory type, and L1, indicating that Spanish speakers' difficulty in identifying the presence of a bilabial or labiodental sound differed depending on whether the auditory word matched or mismatched the written word.

To better understand the three-way interaction, two follow-up fit linear mixed-effects models were run to test for the effect of phoneme type and auditory type independently for the native speakers and Spanish L2 learners of English. These models are shown in Table 4 and Table 5 for the native speakers and Spanish L2 learners, respectively.

| Variable                     | Estimate (SE) | <i>z</i> | <i>P</i> |
|------------------------------|---------------|----------|----------|
| (Intercept)                  | 3.07 (.32)    | 9.75     | <.001    |
| Auditory Type: Mismatch      | -.21 (.23)    | -.73     | >.1      |
| Phoneme Type: Labiodental    | -.002 (.29)   | -.007    | >.1      |
| Auditory Type x Phoneme Type | 1.14 (.48)    | 2.4      | <.05     |

Note: *df* = 1248;  $\alpha$  = .05

Table 4: Logit regression model on native speakers' accuracy results

| Variable                     | Estimate (SE) | <i>z</i> | <i>P</i> |
|------------------------------|---------------|----------|----------|
| (Intercept)                  | 0.46 (.14)    | 3.36     | <.001    |
| Auditory Type: Mismatch      | -.96 (.16)    | -5.95    | <.001    |
| Phoneme Type: Labiodental    | .28 (.16)     | 1.77     | .08      |
| Auditory Type x Phoneme Type | -.11 (.23)    | -.48     | >.1      |

Note: *df* = 1344;  $\alpha$  = .05

Table 5: Logit regression model on Spanish L2 learners' accuracy results

These results showed a clear difference between the two groups. On the one hand, the native speakers (Table 4) showed an interaction between auditory type and phoneme type, indicating that they were statistically more accurate in the mismatch condition, when the auditory stimuli contained a labiodental sound, than in any of the other conditions. On the other hand, the Spanish L2 learners (Table 5) showed a main effect of

auditory type, which indicates that they were more accurate in the match than in the mismatch condition, and a marginal main effect of phoneme type, which indicates that they were marginally more accurate in the labiodental than the bilabial condition.

### **3.6 Discussion**

Experiment 2 was designed to investigate the impact of perception issues on word recognition, by trying to determine if the perception difficulties observed in Experiment 1 affected the learners' ability to recognize words. The results of Experiment 2 showed a distinct division. While the native speakers showed no problems in detecting the target word in the auditory stimuli, the native Spanish speakers had difficulty detecting the target words, particularly in the mismatch conditions. Approached cautiously, these findings imply that the lack of a phonemic distinction in the learners' L1 negatively impacts their ability to access words differentiated by the bilabial-labiodental contrast in English. These word recognition challenges might be connected to the potential occurrence of Spanish L2 learners of English mistakenly activating competing words when they encounter words containing /b/ or /v/ (supported by the results from the mismatch condition).

The concept of L2 learners unintentionally activating competing lexical items has already been documented in the literature that examines L1-L2 category assimilation and word recognition (e.g., Broersma, Cutler, 2011; Pallier, Colomé, Sebastián-Gallés 2001; Weber, Cutler 2004). To perceive spoken words, listeners must match the incoming auditory information with their stored lexical representations. Word recognition models (e.g., Marslen-Wilson, 1987; McClelland, Elman, 1986; Norris, 1994) propose that lexical candidates are activated in multiple parallel pathways and then compete for activation. This means that as the acoustic input unfolds, all words consistent with the input become active in the lexicon until there is enough segmental and suprasegmental information to identify the intended word accurately. When perceiving L2 words, not only words from the L2 but also from the L1 may be activated and compete for recognition. Thus, the fact that L1 does not have a phonemic distinction may result in the activation of words (competitor words) that are not present in the signal, making L2 word recognition less efficient.

Another possibility is that these participants were more sensitive to the written trigger, as they were more accurate whenever they saw a word with a <v>. Interestingly, this pattern was only marginal among the learners, but significant among the native speakers. While still showing high accuracies in all the conditions, native speakers were also more accurate whenever they heard a word with the labiodental sound, but saw a word containing the <b> letter. This interpretation of the results, then, would be in line with previous studies, which have documented the influence of orthography on the pronunciation of [v] (Cartagena, 2002; Stevens, 2000; Takawaki, 2012; Torres Cacoullós, Ferreira, 2000; Trovato, 2017).

In summary, the phonemic inventory of the L1 constrains speech discrimination and the consequent lexical competition they induce (modulated by visual-orthographic information) seems to explain the pattern of results reported in Experiment 2.

#### **4. General discussion and conclusion**

As mentioned before, this study was designed to determine how potential perception issues may affect word recognition in an L2, and how this pattern could be linked to the visual influence of orthography. The results of the study indicate that native speakers of Spanish do indeed show a problem identifying the difference between [b] and [v]. This set of results has already been reported in the literature and the findings have been explained based on the properties of the phonetic inventory of the L1, such that an allophonic contrast in the L1 is not easily identifiable as distinctive phonemes in the L2 (e.g., Abramson, Lisker, 1970; Goto, 1971; Lisker, Abramson, 1967; Polka, 1992; Polka, Werker, 1994; Strange, Akahane-Yamada, Kubo, Trent, Nishi, 2001; Veleva, 1985; Werker, Lalonde, 1988; Werker, Logan, 1985; Werker, Tees, 1984). However, this study goes beyond these findings to determine that these identification problems do hinder word recognition and the results of Experiment 2 can be linked as well to the orthographic properties of English as proposed in the previous discussions. The influence of orthography on the pronunciation of the phoneme [v] has already been reported (Cartagena, 2002; Stevens, 2000; Takawaki, 2012; Torres Cacoullós, Ferreira, 2000; Trovato, 2017) and this study shows that this effect does not only appear in pronunciation but also seems to matter in perception.

In current models of L2 phonetic perception, such as the SLM (Flege, 1995) and PAM (Best, 1995), the perceived similarity between sounds in one's L1 and sounds in the L2 significantly influences the predicted difficulty of different non-native sound distinctions. Other listener-related factors, like the age of acquisition of the L2 and the continued use of the L1, also play a substantial role in determining the performance attained in the L2 (e.g., Flege *et al.*, 1997, 2003). The basis for perceiving similarity varies across these models. Best's PAM model relies on the similarity of articulatory gestures, while Flege's SLM model requires empirical measurement because it cannot be predicted in advance. While this study was not created with the purpose of teasing apart these two models of speech perception, the results do support the claims that the phonetic inventory of the L1 influences the perception of the phonetic inventory in the L2. Future studies on this specific parameter should focus on trying to discriminate between the two models to better understand how the phonemic inventory of the L1 influences the perception of the L2. This study had two main objectives. Firstly, it aimed to investigate the perceptual issues related to allophonic variation in the L1 and its impact on the correct discrimination of minimal pairs in the L2. Secondly, it sought to examine how these perceptual problems affected word recognition in the L2. As of now, this study appears to be one of the first to explore the negative influence of allophonic variation in the L1 on the accurate differentiation of minimal pairs in the L2.

Additionally, just as acoustic prominence can influence the ease of acquiring a sound contrast (see Best, McRoberts, 2003), the visual salience of the contrast is also likely to affect the extent to which learners pay attention to visual cues. Several studies before have a look at how visual information influences auditory perception in an L2. Combining written alphabetic input with auditory input enhances speech perception for different groups of listeners. This has been observed in studies involving individuals initially exposed to the sounds of an unfamiliar language (Erdener, Burnham, 2005) and L2 learners encountering pseudowords (Yazawa *et al.*, 2020). However, if the target contrast is particularly challenging for the listener, this facilitative effect may not be as pronounced (Simon *et al.*, 2010). In this study, both groups showed higher accuracy in identifying sounds when the acoustic information contained a

labiodental sound while visually, they saw the letter "b". This is likely because when they saw the letter <b>, both groups could create a clear mental representation of how this letter should be pronounced, as it is phonemic in both languages. However, when the auditory information did not fully match their mental representation, it was easier for them to recognize that the heard word was not visually present.

In this study, although proficiency did not show a statistically significant influence on the results as previous studies found (e.g., Takawaki, 2012; Tim, 1976; Torres Cacoullos, Ferreria, 2000), it is important to interpret these findings cautiously. The results imply that the impact of L1 phonemic inventory on L2 speech perception might persist across different proficiency levels, even at advanced stages. However, there could be other plausible explanations. For instance, the proficiency (cloze) test used to assess participants' language skills may not have adequately measured their aural proficiency. Since the cloze test relies on the visual modality, it might not have been sensitive enough to capture the variations in L2 learners' auditory perceptual abilities, thereby potentially masking any proficiency-related effects. Future studies should examine the extent to which native-like production, perception, and recognition of /v/ vs. /b/ phonemic contrast is acquirable for Spanish L2 learners whose proficiency is assessed aurally.



**References**

- ABRAMSON, A. S., LISKER, L., Discriminability along the voicing continuum: Cross-language tests, in HÁLA, B., ROMPORTL, M., JANOTA, P. (eds.), *Proceedings of the 6th International Congress of Phonetic Sciences*, Prague, Academia, 1970, 569-573.
- BAAYEN, H. R., *Analyzing Linguistic Data: a Practical Introduction to Statistics Using R*, Cambridge, Cambridge University Press, 2008. <https://doi.org/10.1017/cbo9780511801686>
- BASSETTI, B., Orthography affects second language speech: Double letters and geminate production in English, *JOURNAL OF EXPERIMENTAL PSYCHOLOGY: LEARNING, MEMORY, AND COGNITION*, 2017, **43(11)**, 1835-1842. <https://doi.org/10.1037/xlm0000417>
- BASSETTI, B., HAYES-HARB, R., ESCUDERO, P., Second language phonology at the interface between acoustic and orthographic input, *APPLIED PSYCHOLINGUISTICS*, 2015, **36(1)**, 1-6. <https://doi.org/10.1017/s0142716414000393>
- BASSETTI, B., MAIRANO, P., MASTERSON, J., CERNI, T., Effects of orthographic forms on L2 speech production and phonological awareness, with consideration of speaker level predictors, *LANGUAGE LEARNING*, 2020, **70(4)**, 1218-1256. <https://doi.org/10.1111/lang.12423>.
- BASSETTI, B., SOKOLOVIC-PEROVIC, M., MAIRANO, P., CERNI, T., Orthography-induced length contrasts in the second language phonological systems of L2 speakers of English: Evidence from minimal pairs, *LANGUAGE AND SPEECH*, 2018, **61**, 577-597. <https://doi.org/10.1177/0023830918780141>
- BEST, C. T., A direct realist perspective on cross-language speech perception, in STRANGE, W. (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-language Research*, Timonium, MD, York Press, 1995, 171-204.
- BEST, C. T., MCROBERTS, G. W., Infant perception of non-native consonant contrasts that adults assimilate in different ways, *LANGUAGE AND SPEECH*, 2003, **46(2-3)**, 183-216. <https://doi.org/10.1177/00238309030460020701>
- BROERSMA, M., CUTLER, A., Competition dynamics of second-language listening, *QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY*, 2011, **64(1)**, 74-95. <https://doi.org/10.1080/17470218.2010.499174>
- CARTAGENA, N., *Apuntes para la Historia del Español en Chile*, Santiago, Academia Chilena de la Lengua, 2002.
- ERDENER, V.D., BURNHAM, D.K., The role of audiovisual speech and orthographic information in nonnative speech production, *LANGUAGE LEARNING*, 2005, **55(2)**, 191-228. <https://doi.org/10.1111/j.0023-8333.2005.00303.x>
- ESCUDERO, P., HAYES-HARB, R., MITTERER, H., Novel second-language words and asymmetric lexical access, *JOURNAL OF PHONETICS*, 2008, **36(2)**, 345-360. <https://doi.org/10.1016/j.wocn.2007.11.002>
- ESCUDERO, P., WANROOIJ, K., The effect of L1 orthography on non-native vowel perception, *LANGUAGE AND SPEECH*, 2010, **53(3)**, 343-365. <https://doi.org/10.1177/0023830910371447>

- FLEGE, J.E., Second language speech learning: theory, findings, and problems, *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, Timonium, York Press MD, 1995, 233-277 <https://doi.org/10.1111/j.1600-0404.1995.tb01710.x>
- FLEGE, J.E., BOHN, O.S., JANG, S., Effects of experience on non-native speakers' production and perception of English vowels, *JOURNAL OF PHONETICS*, 1997, **25(4)**, 437-470, Doi: <https://doi.org/10.1006/jpho.1997.0052>
- FLEGE, J.E., SCHILLER, N.O, MEYER, A.S., *Assessing Constraints on Second-language Segmental Production and Perception. Phonetics and Phonology in Language Comprehension and Production: Differences and Similarities*, Berlin, de Gruyter Mouton, 2003, 319-355. <https://doi.org/10.1515/9783110895094.319>
- GOTO, H., Auditory perception by normal Japanese adults of the sounds “ɸ” and “ɹ”, *NEUROPSYCHOLOGIA*, 1971, **9**, 317-323. [https://doi.org/10.1016/0028-3932\(71\)90027-3](https://doi.org/10.1016/0028-3932(71)90027-3)
- HAYES-HARB, R., NICOL, J., BARKER, J., Learning the phonological forms of new words: Effects of orthographic and auditory input, *LANGUAGE AND SPEECH*, 2010, **53(3)**, 367-381. <https://doi.org/10.1177/0023830910371460>
- HOTHORN, T., EVERITT, B. S., *A Handbook of Statistical Analyses Using R*, New York, Chapman & Hall/CRC, 2014. <https://doi.org/10.1201/9781420010657>
- HUALDE, J., *Los sonidos del español: Spanish Language Edition*, Cambridge, Cambridge University Press, 2014. <https://doi.org/10.1017/CBO9780511719943>
- JONGMAN, A., WAYLAND, R., WONG, S., Acoustic characteristics of English fricatives, *THE JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA*, 2000, **108(3)**, 1252-1263. <https://doi.org/10.1121/1.1288413>
- KILBORN, K., MOSS, H., Word Monitoring, *LANGUAGE AND COGNITIVE PROCESSES*, 1996, **11(6)**, 689-694. <https://doi.org/10.1080/016909696387105>
- LISKER, L., ABRAMSON, A.S., The voicing dimension: Some experiments in comparative phonetics, *Haskins Laboratories Status Report on Speech Research*, SR11, 1967, 9-15.
- MARTÍNEZ-GARCÍA, M.T., Syllable Structure Effects in Word Recognition by Spanish- and German-Speaking Second Language Learners of English, *ATLANTIS. JOURNAL OF THE SPANISH ASSOCIATION FOR ANGLO-AMERICAN STUDIES*, 2021, **43(2)**, 1-21. <https://doi.org/10.28914/Atlantis-2021-43.2.01>
- MARSLÉN-WILSON, W. D., Functional parallelism in spoken word-recognition, *COGNITION*, 1987, **25(1-2)**, 71-102. [https://doi.org/10.1016/0010-0277\(87\)90005-9](https://doi.org/10.1016/0010-0277(87)90005-9).
- MATHIEU, L., The influence of foreign scripts on the acquisition of a second language phonological contrast, *SECOND LANGUAGE RESEARCH*, 2016, **32(2)**, 145-170. <https://doi.org/10.1177/0267658315601882>
- MCCLELLAND, J. L., ELMAN, J. L., The TRACE model of speech perception, *COGNITIVE PSYCHOLOGY*, 1986, **18(1)**, 1-86. [https://doi.org/10.1016/0010-0285\(86\)90015-0](https://doi.org/10.1016/0010-0285(86)90015-0)

- NORRIS, D., Shortlist: a connectionist model of continuous speech recognition, *COGNITION*, 1994, **52(3)**, 189–234. [https://doi.org/10.1016/0010-0277\(94\)90043-4](https://doi.org/10.1016/0010-0277(94)90043-4)
- PALLIER, C., COLOMÉ, A., SEBASTIÁN-GALLÉS, N., The influence of native-language phonology on lexical access: exemplar-Based versus abstract lexical entries, *PSYCHOLOGICAL SCIENCE*, 2001, **12(6)**, 445–449. <https://doi.org/10.1111/1467-9280.00383>
- POLKA, L., Characterizing the influence of native experience on adult speech perception, *PERCEPTION AND PSYCHOPHYSICS*, 1992, **52**, 37-52. <https://doi.org/10.3758/bf03206758>
- POLKA, L., WERKER, J. F., Developmental changes in perception of nonnative vowel contrasts, *JOURNAL OF EXPERIMENTAL PSYCHOLOGY: HUMAN PERCEPTION AND PERFORMANCE*, 1994, **20**, 421-435. <https://doi.org/10.1037//0096-1523.20.2.421>
- R DEVELOPMENT CORE TEAM., R: A language and environment for statistical computing, *R foundation for statistical computing*, Vienna, Austria, 2009, 201. <http://www.R-project.org>
- RAFAT, Y., STEVENSON, R. A., Auditory-orthographic integration at the onset of L2 speech acquisition, *LANGUAGE AND SPEECH*, 2019, **62(3)**, 427-451. <https://doi.org/10.1177/0023830918777537>
- SIMON, E., CHAMBLESS, D., ALVES, U.K., Understanding the role of orthography in the acquisition of a non-native vowel contrast, *LANGUAGE SCIENCES*, 2010, **32(3)**, 380-394. <https://doi.org/10.1016/j.langsci.2009.07.001>
- STRANGE, W., AKAHANE-YAMADA, R., KUBO, R., TRENT, S.A., NISHI, K., Effects of consonantal context on perceptual assimilation of American English vowels by Japanese listeners, *JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA*, 2001, **109**, 1691-1704. <https://doi.org/10.1121/1.1353594>
- STEVENS, J., On the Labiodental Pronunciation of Spanish /b/ among Teachers of Spanish as a Second Language, *HISPANIA*, 2000, **83(1)**, 139-149. <https://doi.org/10.2307/346152>
- TAGLIAFERRI, B., *Paradigm*, Perception Research Systems, Inc., 2005. [www.perceptionresearchsystems.com](http://www.perceptionresearchsystems.com)
- TAKAWAKI, S. L., Orthographic loyalty in the Spanish of northern Mexican speakers, *Linguistics and Language Behavior Abstracts (LLBA)*, 2012. Retrieved from: [https://repository.asu.edu/attachments/97661/content//tmp/package5PYacT/Takawaki\\_asu\\_0010N\\_12317.pdf](https://repository.asu.edu/attachments/97661/content//tmp/package5PYacT/Takawaki_asu_0010N_12317.pdf)
- TIM, L., Three consonants in Chicano Spanish: /x/, /b/ and /d/, *BILINGUAL REVIEW / LA REVISTA BILINGÜE*, 1976, **3(2)**, 153-162. Retrieved from: <http://www.jstor.org.lib.utep.edu/stable/25743678>
- TORRES CACOULOS, R., FERREIRA, F., Lexical frequency and voiced labiodental-bilabial variation in new Mexican Spanish, *SOUTHWEST JOURNAL OF LINGUISTICS*, 2000, **19(2)**, 1-17. Retrieved from: <http://search.proquest.com/docview/85527614?accountid=7121>

- TROVATO, A., *A Sociophonetic Analysis of Contact Spanish in the United States: Labiodentalization and Labial Consonant Variation*, UT Electronic Thesis and Dissertations, 2017. <https://repositories.lib.utexas.edu/bitstream/handle/2152/63721/TROVATODISSERTATION-2017.pdf?sequence=1&isAllowed=y>
- VELEVA, M., The perception of the voiced/voiceless phonological opposition in Bulgarian by Arab learners, *SUPOSTAVITELNO EZIKOZNANIE/CONTRASTIVE LINGUISTICS*, 1985, **10**, 28-37.
- WEBER, A., CUTLER, A., Lexical competition in non-native spoken-word recognition, *JOURNAL OF MEMORY AND LANGUAGE*, 2004, **50(1)**, 1–25. [https://doi.org/10.1016/s0749-596x\(03\)00105-0](https://doi.org/10.1016/s0749-596x(03)00105-0)
- WERKER, J.F., LALONDE, C.E., Cross-language speech perception: Initial capabilities and developmental change, *DEVELOPMENTAL PSYCHOLOGY*, 1988, **24**, 672-683. <https://doi.org/10.1037/0012-1649.24.5.672>
- WERKER, J.F., LOGAN, J.S., Cross-language evidence for three factors in speech perception, *PERCEPTION & PSYCHOPHYSICS*, 1985, **37**, 35-44. <https://doi.org/10.3758/bf03207136>
- WERKER, J.F., TEES, R.C., Cross-language speech perception: Evidence for perceptual reorganization during the first year of life, *INFANT BEHAVIOR AND DEVELOPMENT*, 1984, **7**, 49-63. [https://doi.org/10.1016/s0163-6383\(02\)00093-0](https://doi.org/10.1016/s0163-6383(02)00093-0)
- YAZAWA, K., WHANG, J., KONDO, M., ESCUDERO, P., Language-dependent cue weighting: An investigation of perception modes in L2 learning, *SECOND LANGUAGE RESEARCH*, 2020, **36(4)**, 557-581. <https://doi.org/10.1177/0267658319832645>

**María Teresa MARTÍNEZ GARCÍA** is an assistant professor in the Department of Languages at the University of Utah, Asia campus. Her scientific specialization focuses on speech perception in a second language (L2) and bilingualism. In addition, she has experience in the field of teaching Spanish as an L2 and years of experience teaching Spanish, English, and Linguistics in the United States, Spain, and South Korea.

**Appendix A: List of experimental stimuli in Experiment 2**

| Forms with /b/  | Forms with /v/  |
|---|---|
| She said the word "bail" during her speech.                                 | She said the word "veil" during her speech.                                 |
| The water was bailed, so that the passengers could not suspect the problem. | The water was veiled, so that the passengers could not suspect the problem. |
| "Bailing" is the word of the day.   | "Veiling" is the word of the day.   |
| His brother bails him with he is in trouble.                                | His brother veils him with he is in trouble.                                |
| He bought a new balance for the kitchen.                                    | He bought a new valance for the kitchen.                                    |
| She looked out across the field and saw several bales.                      | She looked out across the field and saw several vales.                      |
| Paying for the ballet is a luxury we cannot afford right now.               | Paying for the vallet is a luxury we cannot afford right now.               |
| The candidate wanted to buy some boats.                                     | The candidate wanted to buy some votes.                                     |
| Write "banish" in the blank.  | Write "vanish" in the blank.  |
| He doesn't know how to say "banished".                                      | He doesn't know how to say "vanished".                                      |
| He feels that "banishes" is hard to pronounce.                              | He feels that "vanishes" is hard to pronounce.                              |
| Joan saw "banishing" written on the page.                                   | Joan saw "vanishing" written on the page.                                   |
| That bat is heavier than it looks.  | That vat is heavier than it looks.  |
| There are seven bats in the attic.  | There are seven vats in the attic.  |
| This machine was invented to bend things.                                   | This machine was invented to vend things.                                   |
| Please use "bended" in a sentence.  | Please use "vended" in a sentence.  |
| This is a bending machine.  | This is a vending machine.  |
| At his job, the employee bends things.                                      | At his job, the employee vends things.                                      |
| I gave her my best.   | I gave her my vest.   |
| This bet has become rather famous.  | This vet has become rather famous.  |
| I wore a biking costume.  | I wore a viking costume.  |
| We often boat in the summer.  | We often vote in the summer.  |
| I have always liked boating.  | I have always liked voting.   |
| It is difficult to measure one bolt.  | It is difficult to measure one volt.  |
| The peaseants will bow to those in power.                                   | The peasants will vow to those in power.                                    |
| "Bowel" is the last word he looked up.                                      | "Vowel" is the last word he looked up.                                      |
| This is a berry flavored pie.   | This is a very flavored pie.  |
| The knight bowed to the king.   | The knight vowed to the king.   |
| That curb is treacherous!   | That curve is treacherous!  |

|   |   |
|---|---|
| Mary always trips on the curbed sidewalk.                         | Mary always trips on the curved sidewalk.                         |
| Anne believes that curbing makes the street more dangerous.       | Anne believes that curving makes the street more dangerous.       |
| There are several curbs along the way.                            | There are several curves along the way.                           |
| I won't listen to you dribble any longer.                         | I won't listen to you drivel any longer.                          |
| The password is "dub".  | The password is "dove".   |
| These dubs are worth a lot of money.                              | These doves are worth a lot of money.                             |
| The woman found a fibre in her bag.                               | The woman found a fiver in her bag.                               |
| Please say the word "fibres" as carefully as you can.             | Please say the word "fivers" as carefully as you can.             |
| The sound of the gabble was overwhelming.                         | The sound of the gavel was overwhelming.                          |
| The lobes were divided into smaller pieces.                       | The loaves were divided into smaller pieces.                      |
| This is quite the marble!   | This is quite the marvel!   |
| The marbled cake was prepared for the birthday.                   | The marvelled cake was prepared for the birthday.                 |
| The collector keeps her marbles in a temperature-controlled room. | The collector keeps her marvels in a temperature-controlled room. |
| The king was not impressed with the rebel.                        | The king was not impressed with the revel.                        |
| Robin Hood was pleased with the rebels in the crowd.              | Robin Hood was pleased with the revels in the crowd.              |
| I will say the word "robe" now.                                   | I will say the word "rove" now.                                   |
| Please say "robed" now.   | Please say "roved" now.   |
| The child missed "robes" on the spelling test.                    | The child missed "roves" on the spelling test.                    |
| I never know how to use "robing" in a sentence.                   | I never know how to use "roving" in a sentence.                   |