UNIVERSITAT DE BARCELONA

Estudios de Fonética Experimental
Journal of Experimental Phonetics
ISSN: 1575-5533 · ISSN-e: 2385-3573

EFE Estudios de Fonética Experimental
Journal of Experimental Phonetics

Orthographic influence on English word recognition by Spanish and Korean learners

Maria Teresa Martinez Garcia [1] 0000-0003-3187-0853

[1] Universidad de Valladolid (Spain)

DOI: 10.1344/efe-2025-34-113-128

Corresponding address: mariateresamg@uva.es

Received: 23-08-2024 Accepted: 06-03-2025 Published: 16-06-2025

Martinez Garcia, M. T. (2025). Orthographic influence on English word recognition by Spanish and Korean learners. *Estudios de Fonética Experimental*, *34*, 113–128. https://doi.org/10.1344/efe-2025-34-113-128

Abstract

This study explores how visual-orthographic information influences the perception and recognition of phonemes in a second language (L2). Spanish and Korean learners of English participated in tasks focusing on the English labiodental /v/ and bilabial /b/ phonemes, which are either allophonic or non-existent in their native languages but phonemic in English. The findings demonstrate that the participants' first language (L1) significantly impacts their ability to perceive and recognize these English phonemes. Additionally, orthographic cues were found to play a crucial role in this process. The results support models that emphasize the importance of L1-L2 phoneme mapping and highlight the influence of orthographic information in L2 word recognition.

Keywords

L2 mapping; orthographic information; word recognition; speech perception

La influència de l'ortografia en el reconeixement de mots en anglès per part d'aprenents parlants d'espanyol i de coreà

RESUM

Aquest estudi explora com la informació visual-ortogràfica influeix en la percepció i el reconeixement de fonemes en una segona llengua (L2). Diversos aprenents d'anglès de parla espanyola i coreana van participar en tasques centrades en els fonemes labiodental /v/ i bilabial /b/ de l'anglès, que són al·lofònics o inexistents en les seves llengües maternes però fonèmics en anglès. Els resultats demostren que la llengua materna dels participants (L1) té un impacte significatiu en la seva capacitat per percebre i reconèixer aquests fonemes en anglès. A més, es troba que les pistes ortogràfiques exerceixen un paper crucial en aquest procés. Els resultats donen suport a models que subratllen la importància del mapatge de fonemes entre L1 i L2, i ressalten la influència de la informació ortogràfica en el reconeixement de paraules en L2.

MOTS CLAU

mapatge L2; informació ortogràfica; reconeixement de paraules; percepció de la parla

La influencia de la ortografía en el reconocimiento de palabras en inglés por aprendices hispanohablantes y coreanos

RESUMEN

Este estudio explora cómo la información visual-ortográfica influye en la percepción y el reconocimiento de fonemas en una segunda lengua (L2). Estudiantes de inglés de habla hispana y coreana participaron en tareas centradas en los fonemas labiodental /v/ y bilabial /b/ del inglés, que son alofónicos o inexistentes en sus lenguas maternas, pero fonémicos en inglés. Los resultados demuestran que la lengua materna de los participantes (L1) impacta significativamente en su capacidad para percibir y reconocer estos fonemas en inglés. Además, se encontró que las pistas ortográficas desempeñan un papel crucial en este proceso. Los resultados apoyan modelos que enfatizan la importancia del mapeo de fonemas entre la L1 y la L2, y destacan la influencia de la información ortográfica en el reconocimiento de palabras en L2.

PALABRAS CLAVE

mapeo L2; información ortográfica; reconocimiento de palabras; percepción del habla

1. Introduction

Across the world, there are numerous alphabets and writing systems, each with unique features that influence how languages are processed and perceived. From the Latin alphabet used in English and Spanish, to the Cyrillic script in Russian, and logographic systems like Chinese, these orthographic systems provide distinct visual representations of spoken language. These scripts not only encode linguistic information but can also play a crucial role in how speakers of a language perceive and produce sounds (e.g., Bassetti, 2017; Bassetti et al., 2015, 2018, 2020). The interaction between spoken and written language is particularly relevant in second language (L2) learning, where learners often rely on familiar orthographic conventions to navigate new phonetic landscapes (e.g., Erdener & Burnham, 2005; Hayes-Harb et al., 2010; Escudero & Wanrooij, 2010). For instance, speakers of languages with alphabetic scripts may experience different challenges in L2 speech perception and production compared to those whose native languages use syllabaries or logograms (e.g., Mok et al., 2018; Cutler, 2015; Mathieu, 2016). This study explores how the diverse orthographic systems across languages can shape phonological perception, especially in the context of learning English as an L2.

Recent research has increasingly focused on the impact of orthographic forms (spelling conventions) on aspects of L2 speech processing, including perception, production, and sound categorization (e.g., Bassetti, 2017; Hayes-Harb et al., 2010; Escudero et al., 2008). Most studies have focused on languages that share the same alphabet, particularly those using the Latin script (e.g., Bassetti et al., 2018; Escudero & Wanrooij, 2010). However, the effect of orthography on speech perception in languages with non-traditional or non-alphabetic scripts is an emerging area of interest. For example, research on Chinese, a language with a logographic writing system, has shown that orthography can influence the perception of tone and segmental contrasts in L2 learners (e.g., Mok et al., 2018). Similarly, studies on Japanese kana, a syllabary script, have shown that L2 learners might transfer syllable-based orthographic patterns from their L1 to their L2, affecting their perception and production of foreign sounds (e.g., Cutler, 2015). These findings suggest that the influence of orthography extends beyond alphabetic systems, affecting L2 learners from diverse linguistic backgrounds in distinct ways. Understanding how different scripts interact with speech perception and production in L2 contexts is crucial for developing comprehensive models of L2 phonological acquisition.

Building on this growing interest, recent studies have delved deeper into the role of orthographic forms in shaping specific aspects of L2 speech processing, such as perception, production, and sound categorization. Research by Bassetti and colleagues, among others, has shown that L2 speakers may produce sound contrasts not present in the target language, a phenomenon influenced by the prominence of visual-orthographic cues (e.g., Bassetti, 2017; Bassetti et al., 2015, 2018, 2020). However, most of the previous studies, as it will be reviewed in this section, have focused on languages sharing the same alphabet. This study examines how challenges in speech perception and sound categorization might affect word recognition in Spanish- and Korean-speaking learners of English as an L2, with a focus on the role of orthographic information in modulating this effect, and contrasting two native languages (L1) with different writing systems.

In both English and Spanish orthography, the graphemes <b> and <v> are present, but they are only phonetically distinguished in English. In Spanish, both graphemes are typically pronounced as [b] or [β̞], depending on phonetic context and regional variation[1], whereas in English, <v> is realized as a labiodental fricative and <b> as a bilabial plosive (Hualde, 2014; Jogman, Wayland, & Wong, 2000).

[1] In some regions of Spain, particularly those influenced by Catalan, a distinction between /b/ and /v/ may be observed. This distinction is less common in Latin American varieties.

The standard Spanish phonemic inventory lacks the voiced labiodental fricative /v/, although Hualde (2014) notes its occurrence in specific contexts as a result of coarticulation, such as in the word *afgano* (Afghan), pronounced [avˈɣ̞ano]. This difference between the phonemic systems of English and Spanish, despite their similar orthographic representations, is central to this study. In contrast, the Korean phonemic inventory lacks labial fricatives and includes only labial stops: /p/, /ph/, /p', and [b] (Shin et al., 2012). The English /f/ and /v/ are perceptually assimilated to the Korean lenis stop /p/ when borrowed into Korean, as part of cross-language perceptual assimilation in loanword phonology (e.g., Tranter, 2000). For example, the English word *video* is adapted as [pitio] and *coffee* as [kʰopi]. This process reflects cross-language perceptual assimilation rather than a phonological adaptation process, as learners map non-native sounds to the closest available categories in their native phonemic inventory.

The occurrence of the voiced labiodental fricative [v] in the speech of Spanish-English bilinguals, particularly among Spanish-speaking adults learning English as an L2, has been extensively examined in previous research. One perspective suggests that exposure to the English phoneme /v/ affects the Spanish sound system, leading to the emergence of [v] as a variant of the Spanish phoneme /b/. This indicates that English language proficiency and exposure can influence how Spanish speakers produce and distinguish sound contrasts in their native language (e.g., Takawaki, 2012; Trovato, 2017). This phenomenon is especially pronounced in bilingual communities, such as those in New Mexico and California, where significant language contact occurs. The degree of influence often correlates with the individual's English proficiency level (e.g., Takawaki, 2012; Timm, 1976; Torres Cacoullos & Ferreira, 2000). Additionally, Stevens (2000) observed that Instructors who are native Spanish speakers teaching English as a second language frequently produce [v], attributing this to the influence of English orthography. Moreover, Martínez-García (2023) found that the presence of orthographic information can modulate the perception and word recognition of English phonemes by Spanish-speaking learners, particularly when distinguishing between labiodental /v/ and bilabial /b/ phonemes. These findings highlight the importance of orthographic influence on phonetic adaptation in bilinguals, particularly in contexts of high language contact and growing L2 proficiency. However, studies focusing on this specific aspect in Korean learners of English are scarce, indicating a need for further research in this area.

In addition to language contact and proficiency, orthographic forms significantly influence speech perception in both native speakers and L2 learners. The literature demonstrates that visual-orthographic input can either aid or hinder speech perception. In some cases, alphabetic orthographic cues, when combined with auditory input, enhance the perception of speech sounds (e.g., Erdener & Burnham, 2005; Escudero et al., 2008). However, discrepancies between spoken and written forms can lead to misperceptions (e.g., Hayes-Harb et al., 2010; Escudero & Wanrooij, 2010, Experiment 2; Mathieu, 2016; Rafat & Stevenson, 2019). These misperceptions often manifest as sound additions, omissions, or substitutions during the perception process. Additionally, research on L2 speech production supports these findings, revealing that orthographic forms can induce sound changes in L2 pronunciation (for a review, see Bassetti et al., 2015). Despite the established influence of orthography on speech perception, production, and metalinguistic awareness, there remains a notable gap in research on its effect on word recognition in L2 learners.

This study addresses the gap in understanding how L2 learners acquire non-native phonemic contrasts by investigating the potential impact of perception issues on word recognition, with particular attention to the visual influence of orthography. By including Korean, the study enables a comprehensive comparison of how orthography influences phonetic distinctions across languages. In English, orthography maintains distinct phonemes like /b/ and /v/; in Spanish, orthographic distinctions exist without phonetic contrast; and in Korean, certain fricatives are not represented at all. This comparison highlights how orthographic systems shape phonological perception, loanword adaptation, and the maintenance or neutralization of phonetic distinctions, offering valuable insights into the interplay between written and spoken language. The study follows the same methodology as that described in Martinez-Garcia (2023), ensuring consistency in the approach to analyzing orthographic influence across different linguistic contexts.

2. Experiment 1: AXB task

2.1. Participants

32 native speakers of English (16 females; mean age=23 years) from a midwestern university in the USA, 32 Spanish-speaking learners (12 females; mean age=24 years) from the University of Alicante (Spain), and 29 Korean-speaking learners (20 female; mean age=22 years) from the University of Utah Asia Campus (Incheon, South Korea) completed this study. The two learners' groups were matched in different individual differences measures, as can be seen in Table 1, where the proficiency was established by a cloze test (Brown, 1980) out of 50 possible points. None of the participants reported any visual or hearing impairment.

	Spanish learners	**Korean learners**
Age of acquisition	12 (2.3)	11 (2.3)
Years of L2 instruction	12 (3.6)	11 (4.8)
Months living in an L2 environment	15 (36.8)	25 (31.17)
English proficiency	25.94 (12.04)	28 (8.52)

Table 1. Average and standard deviation individual data of the two groups.

Participants in the Spanish-speaking group were from the Valencia-speaking region of Spain. While this region is influenced by Catalan, all participants reported being Spanish-dominant speakers. Therefore, any potential influence of Catalan phonology on their perception or production of /b/ and /v/ was considered unlikely but cannot be entirely ruled out. Participants came from diverse academic backgrounds unrelated to linguistics or language studies, which suggests that their responses were less likely to be influenced by formal training in phonetics or linguistics.

2.2. Materials

Participants completed an AXB discrimination task using minimal pairs of eighteen nonce words that contrasted the English phonemes /b/ and /v/.

Each nonword pair was presented in four possible trial orders (AAB, ABB, BAA, BAB), resulting in 72 distinct trials. To minimize order effects and ensure balanced exposure, the presentation of trials was counterbalanced among participants, and individual participants did not hear the same item more than once.

The nonwords were carefully constructed to conform to English phonotactic rules and were verified to ensure they did not correspond to real English words. The stimuli were recorded by three native English speakers with different regional accents, and their roles in the trial sequences (A, X, B) were rotated across trials to reduce speaker-related biases. See Appendix A.

To prevent participants from focusing exclusively on the /b/-/v/ distinction, 72 filler trials were included. These fillers were divided into two conditions: (1) contrasts involving the presence or absence of a segmental unit (e.g., schwa), and (2) contrasts differing by a single phoneme. These fillers were designed to divert attention from the target contrast and maintain participant engagement without contributing to the primary research question.

Before the main session, participants completed six practice trials with feedback to ensure task comprehension. During the main session, participants were instructed to identify whether the middle nonword (X) was more similar to the first (A) or the third (B) using a mouse click. Feedback was not provided during the experimental trials.

2.3. Procedures

Paradigm software developed by Perception Research Systems, Inc (Tagliaferri, 2005), was used to present the stimuli. Participants, seated comfortably at approximately 20–30 cm from the screen in a quiet room, received instructions in their native language to carefully listen to a sequence of three nonce words and determine whether the second word (X) was more similar to the first or the third word (A or B). All visual information was displayed in Arial font at 20 points. The inter-stimulus interval was set at 1,000 ms, and participants made their selection by clicking one of two mouse buttons. The subsequent trial started immediately after the participant click on the mouse to respond to the previous trial. The inter-stimulus interval (ISI) of 1,000 ms was chosen to allow participants sufficient time to process each stimulus fully, reducing cognitive load for the subtle /b/ and /v/ contrast.

Before the main session, which did not include feedback, participants completed a practice session consisting of six stimuli, during which feedback was provided. All trials were fully randomized for each participant. The task took approximately 20–25 minutes to complete, including practice trials.

2.4. Data analysis

The accuracy of participants was examined using a logistic regression model, following the approach outlined by Baayen (2008). This analysis was conducted with the *glm* package (Hothorn & Everitt, 2014) in *R* (R Core Team, 2009). The participants' first language (L1) was treated as a categorical predictor with three levels (English, Spanish, and Korean), with the English group serving as the baseline. Phoneme type was also treated as a categorical predictor with two levels (Bilabial vs. Labiodental), with the bilabial phoneme as the baseline, given that this phoneme is present in all the participants' L1s. The influence of these predictors was evaluated using log-likelihood tests, comparing models with and without the predictors.

Two sets of models were developed: one focused on the overall accuracy rates using L1 and phoneme type as predictors (and their interaction), and another focused on the L2 learners' accuracy rates, incorporating phoneme type, proficiency, and their interaction. The impact of L2 proficiency was evaluated by comparing models that included proficiency against those that did not, with the best-fitting model retained. Since the inclusion of proficiency did not enhance the model, only the analysis of overall accuracy rates based on L1 and phoneme type (excluding interaction) is presented. Participants and items were included as random variables in the analysis.

2.5. Results

Figure 1 shows the average accuracy results for the three groups across the two phoneme conditions, while Table 2 details the outcomes of the logit mixed-effects model for the accuracy of all participants.

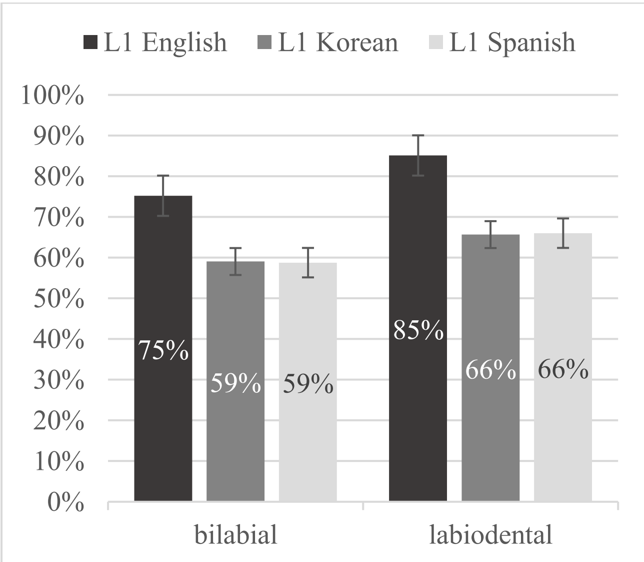

Figure 1. Mean accuracy (standard errors) of the three groups in the AXB task.

Variable	B	SE	z	p
(Intercept)	1.27	.14	8.35	<.001
L1: Spanish L2 learners	-0.94	.17	-5.45	<.001
L1: Korean L2 learners	-0.94	.17	-5.48	<.001
Phoneme Type: Labiodent.	0.39	.12	3.32	<.001

Note: $df = 1494$; $\alpha = .05$

Table 2. Logit regression model on all participants' accuracy results.

2.6. Discussion

In this initial experiment, an AXB task was used to investigate whether Spanish and Korean L2 learners of English would encounter difficulties in discriminating between the bilabial and labiodental contrasts in English, a sound contrast that is either allophonic (in Spanish) or non-existent (in Korean) in their L1s. The results showed that native speakers outperformed L2 learners in accuracy, and that the labiodental sound was more accurately perceived than the bilabial sound. However, the absence of an interaction between group and phoneme type indicates that, despite native speakers being more accurate overall, both Spanish and Korean L2 learners exhibited the same pattern of results, regardless of whether they were listening to a bilabial sound (which exists in their L1) or a labiodental sound (which does not exist in their L1 as a phoneme).

These discrimination results align with the phonemic inventories of the learners' native languages. Specifically, the Spanish speakers' perception reflected the allophonic nature of the [b] versus [v] distinction in their L1. Given that these sounds are allophonic in Spanish, albeit in very limited contexts, Spanish L2 learners of English experienced difficulties in discriminating between them. The Korean learners, on the other hand, struggled with both sounds due to the absence of labiodental fricatives in Korean, highlighting the challenge of acquiring not only a novel phonemic contrast but also the individual sounds themselves, which are absent in their native language.

The difficulties faced by both groups are consistent with established research indicating that the absence of a phonemic distinction in the L1 can significantly hinder L2 perception. When L2 learners are exposed to phonemic contrasts that do not exist in their native language, they may fail to develop the perceptual categories necessary to distinguish these sounds accurately (e.g., Best, 1995; Flege, 1995; Kuhl, 1991). This often results in learners assimilating the novel L2 sounds to the closest native categories, leading to persistent difficulties in perception and production. For instance, Japanese learners of English notoriously struggle to differentiate between /l/ and /r/ due to the lack of this distinction in Japanese (Goto, 1971; Lively et al., 1994), and French learners face challenges with English /h/ as it does not exist in French phonology (Caramazza et al., 1973; Melnik & Peperkamp, 2021).

Separate scores for /b/ and /v/ were analyzed to explore potential asymmetries in the perception of these phonemes. This approach allowed us to examine whether acoustic salience contributed to differences in accuracy, as /v/ is often reported to be more salient than /b/ in perception studies. The lower-than-ceiling performance of native speakers in the AXB task may be attributed to the use of nonwords, which removed lexical cues and heightened task difficulty. Previous research has shown that even native speakers can experience reduced accuracy when processing fine-grained phonemic contrasts without lexical context (e.g., Escudero et al., 2008).

These findings contribute to the existing evidence that contrasts which are allophonic or non-existent in the L1 pose challenges when they are phonemic in the L2 (e.g., Abramson & Lisker, 1970; Best, 1995; Flege, 1995; Polka, 1992; Polka & Werker, 1994; Strange et al., 2001; Werker & Tees, 1984). The perceptual difficulties observed in this experiment suggest that the lack of phonemic distinctions in the L1 can lead to misperceptions in the L2, as learners may not be attuned to the subtle acoustic cues that differentiate similar sounds in the target language (e.g., Werker & Logan, 1985; Werker & Lalonde, 1988). These misperceptions can have broader implications for L2 acquisition, particularly in the domain of word recognition.

Previous studies suggest that difficulties in accurately perceiving L2 sounds can increase lexical competition, thereby reducing the efficiency of word recognition (e.g., Broersma & Cutler, 2011; Escudero, 2007; Weber & Cutler, 2004). When learners misperceive phonemic contrasts, they may incorrectly activate multiple lexical candidates or "phantom words", leading to confusion and slower recognition times (Broersma & Cutler, 2008). For example, Spanish or Korean speakers might struggle to distinguish between "bat" and "vat" in English, resulting in increased difficulty in recognizing these words in continuous speech. Consequently, Experiment 2 was designed to explore how misperception of this phonemic contrast in English might influence L2 word recognition and to examine the role of orthographic information in perceiving this sound contrast, specifically looking at two very distinct scripts.

3. Experiment 2: Word monitoring task

3.1. Participants

In Experiment 2, the participants from Experiment 1 were also included. To avoid any potential impact on their word recognition abilities, they completed the word monitoring task prior to the AXB task, ensuring that their performance was not influenced by patterns or details they might have observed during Experiment 1. Participants completed both experiments consecutively, with no significant breaks between them.

3.2. Materials

In this experiment, all groups participated in a word monitoring task, where they were required to detect a pre-specified target word within the auditory input, following the methodology outlined by Kilborn and Moss (1996). Participants were asked to monitor for a target word that either included or excluded a specific phoneme (e.g., "fiber" vs. "fiver") within semantically ambiguous sentences that could either align with or differ from the target word (i.e., match or mismatch the phonetic content). An example of the experimental items is outlined in Table 3.

Condition	Written	Aud. expected	Aud. actual
Match bilabial	*FIBER*	/b/	/b/
Match labiodental	*FIVER*	/v/	/v/
Mismatch bilabial	*FIBER*	/b/	/v/
Mismatch labiodental	*FIVER*	/v/	/b/

Table 3. Stimuli examples of the word monitoring task (e.g., "The woman found a fiber/fiver in her bag").

The experiment included 48 target items, each containing either the phoneme /b/ or /v/ in various word positions (e.g., "best" vs. "vest"). To ensure participants could not rely solely on lexical cues, sentences were designed to be semantically ambiguous. Two native English speakers assessed these sentences for both plausibility and ambiguity. Additionally, the target word's position within the sentence was varied to avoid creating any participant expectations.

The experiment also included 96 filler items, designed similarly to those used in Experiment 1. Half of these fillers consisted of 36 minimal and 12 near-minimal pairs, where a vowel was added or omitted at the start of /s/-initial clusters (e.g., "especial" vs. "special"). The other half of the fillers varied either by the number of phonemes or by a single phoneme (e.g., "bead" vs. "bee" or "swing" vs. "wing"). These phonetic contrasts occurred at different positions within the words, including word-initial, word-medial, and word-final positions. Fillers were included to divert participants' attention from the specific phoneme contrasts being tested, minimizing task bias and ensuring engagement throughout the task. All sentences, both experimental and filler, were recorded by a female native speaker of American English with a Midwestern accent (Speaker 3 from Experiment 1). The stimuli were presented using a randomized Latin square design.

The target words in the Word Monitoring Task were selected based on their frequency and common usage in English, aligning with participants' expected proficiency levels. A formal lexical familiarity test was administered, yielding no difference among the groups and suggesting that participants were likely to have established phonological representations of them. See Appendix B.

3.3. Procedures

The stimuli were presented using Paradigm software. In each trial, participants were shown the target word in capital letters at the center of the screen for 1,000 ms (e.g., FIBER). As the word disappeared, the corresponding audio played. Participants then listened to a sentence that either included or did not include the target word they had just seen (e.g., "The woman found a fiber in her bag." or "The woman found a fiver in her bag."). Their task was to determine whether the sentence contained the displayed word by clicking a button labeled with the equivalent of 'YES' or 'NO' in their respective native language, using a mouse. Participants could respond as soon as they recognized the word in the sentence, even if the audio had not finished playing, or they could wait until the end of the sentence to confirm the word's absence. Once a response was made, the next trial began immediately.

The experiment started with six practice trials that included feedback. However, during the main session, participants did not receive any feedback.

3.4. Data analysis

Participants' accuracy was analyzed using a logistic regression model, similar to the approach used in Experiment 1. Three categorical predictors were included in the analysis. The first predictor was L1, with three levels (English, Spanish, and Korean), using English as the reference category. The second predictor was the type of phoneme (e.g., /b/ vs. /v/), with the bilabial sound serving as the reference. Lastly, the match between the monitored word and the auditory stimulus (match vs. mismatch) was examined, with "Match" as the baseline category.

Two sets of models were run for this analysis. The first set used L1 as a predictor and considered accuracy across all participants. The second set focused on L2 learners, examining their accuracy in relation to proficiency. However, proficiency was ultimately excluded from the final model, as the model without it provided the best fit. Both sets included participant and item as random effects.

3.5. Results

Figure 2 shows mean accuracy in the word monitoring task, divided by auditory type (match vs. mismatch) and phoneme type (bilabial vs. labiodental). Table 4 reports the logistic regression results for overall accuracy.

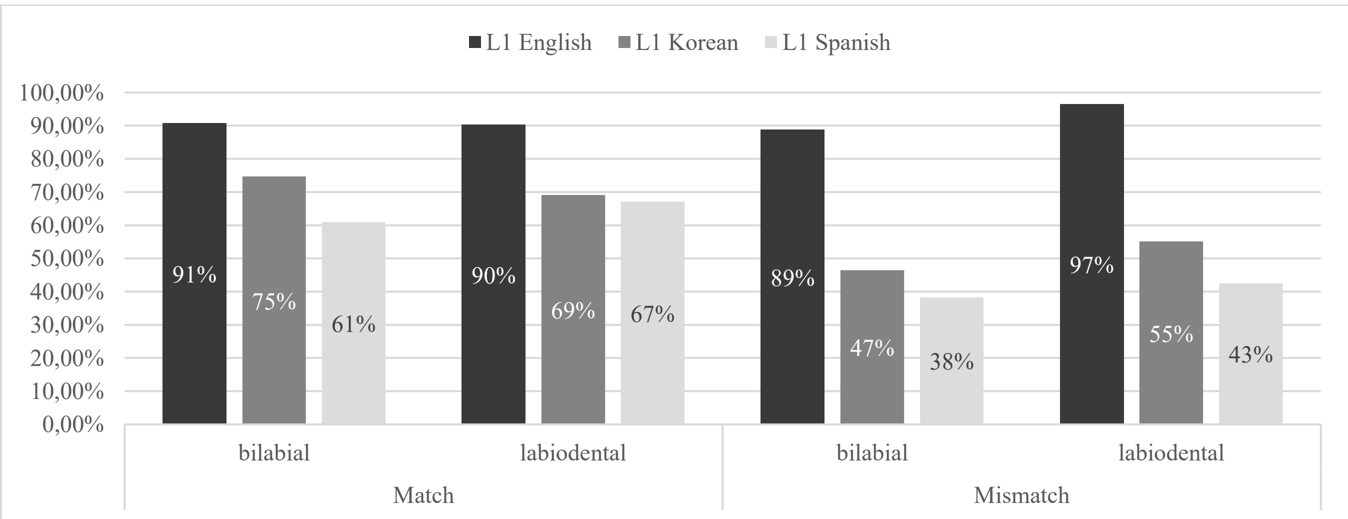

Figure 2. Mean accuracy (SD) of the three L1 groups in the word monitoring task. Bilabial and labiodental labels refer to the auditory stimuli.

Variable	B	SE	z	p
(Intercept)	2.38	.22	11.07	<.001
L1: Spanish L2 learners	-1.91	.26	-7.4	<.001
L1: Korean L2 learners	-1.23	.27	-4.6	<.001
Auditory type: Mismatch	-0.22	.26	-.85	>.1
Phoneme type: Labiodent.	-0.05	.27	-0.18	>.1
L1 Spanish × Auditory t.	-0.75	.31	-2.4	<.05
L1 Korean × Auditory t.	-1.08	.32	-3.4	<.001
L1 Spanish × Phoneme t.	0.33	.31	1.07	>.1
L1 Korean × Phoneme t.	-0.25	.32	-0.77	>.1
Auditory t. × Phoneme t.	1.34	.46	2.94	<.01
L1 Spanish × Auditory t. × Phoneme t.	-1.45	.51	-2.84	<.01
L1 Korean × Auditory t. × Phoneme t.	-0.68	.52	-1.32	>.1

Note: $df = 3839$; $\alpha = .05$

Table 4. Logit regression model on all participants' accuracy results.

The model results revealed several key effects: a main effect of L1, indicating that L2 learners (both Spanish and Korean) exhibited different accuracy patterns compared to native speakers; an interaction between auditory type and L1 for Spanish and Korean learners, showing that, unlike native speakers, these learners had varying accuracy when identifying words with /b/ versus /v/; an interaction between auditory type and phoneme type, suggesting that the accuracy of identifying bilabial versus labiodental sounds depended on whether the word was in the match or mismatch condition; and a three-way interaction between phoneme type, auditory type, and L1, highlighting that Spanish speakers (but not Korean speakers) experienced varying difficulty in recognizing bilabial or labiodental sounds based on whether the auditory word matched or mismatched the written word.

To further explore the three-way interaction, three additional linear mixed-effects models were conducted to examine the effects of phoneme type and auditory type separately for each of the three groups. The results of these models are presented in Table 5 for native speakers, Table 6 for Spanish L2 learners, and Table 7 for Korean L2 learners.

Variable	B	SE	z	p
(Intercept)	3.07	.32	9.75	<.001
Auditory type: Mismatch	-.21	.23	-.73	>.1
Phoneme type: Labiodent.	-.002	.29	-.007	>.1
Auditory t. × Phoneme t.	1.14	.48	2.4	<.05

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

Table 5. Logit regression model on native speakers' accuracy results.

Variable	B	SE	z	p
(Intercept)	0.46	.14	3.36	<.001
Auditory type: Mismatch	-.96	.16	-5.95	<.001
Phoneme type: Labiodent.	.28	.16	1.77	.08
Auditory t. × Phoneme t.	-.11	.23	-.48	>.1

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

Table 6. Logit regression model on Spanish L2 learners' accuracy results.

Variable	B	SE	z	p
(Intercept)	1.16	.16	7.06	<.001
Auditory type: Mismatch	-1.32	.18	-7.22	<.001
Phoneme type: Labiodent.	-0.29	.18	-1.64	>.1
Auditory t. × Phoneme t.	0.67	.25	2.67	<.001

Note: $df = 1247$; $\alpha = .05$

Table 7. Logit regression model on Korean L2 learners' accuracy results.

The results highlighted a distinct difference between the groups. Native speakers (Table 5) showed a main effect of auditory type, indicating that they were better in the match than in the mistmatch condition and an interaction between auditory type and phoneme type, demonstrating statistically higher accuracy in the mismatch condition when the auditory stimuli featured a labiodental sound compared to the other conditions. In contrast, the Spanish L2 learners (Table 6) showed a main effect of auditory type, indicating higher accuracy in the match condition than in the mismatch condition, along with a marginal main effect of phoneme type, suggesting slightly better accuracy in the labiodental condition compared to the bilabial condition. Similarly to the native speakers' pattern, the Korean L2 learners (Table 7) showed a main effect of auditory type, indicating that they were better in the match than in the mismatch condition and an interaction between auditory type and phoneme type, demonstrating statistically higher accuracy in the mismatch condition when the auditory stimuli featured a bilabial sound compared to the other conditions.

3.6. Discussion

Experiment 2 aimed to explore how perceptual difficulties affect word recognition by examining whether the challenges observed in Experiment 1 influenced learners' ability to identify words and how much of an effect the orthographic representation of words could have on this word recognition ability. The results revealed a clear distinction between the groups. Native speakers had no difficulty detecting the target words in the auditory stimuli, whereas native Spanish and Korean speakers struggled, particularly in the mismatch conditions. These findings suggest that the absence of a phonemic distinction in the learners' L1 hampers their ability to recognize words that differ by the bilabial-labiodental contrast in English. However, while Spanish L2 learners of English did only show a marginal interaction between the auditory and phoneme types, both native speakers and Korean L2 learners showed an interaction, which indicated that the effect of the fricative phoneme was larger in the mismatch than in the match condition.

The word recognition issues observed may be related to L2 learners mistakenly activating competing words when encountering words with /b/ or /v/, as indicated by the mismatch condition results. This phenomenon, where L2 learners inadvertently activate competing lexical items, has been documented in studies on L1-L2 category assimilation and word recognition (e.g., Broersma & Cutler, 2011; Pallier et al., 2001; Weber & Cutler, 2004). Word recognition models (e.g., Marslen-Wilson, 1987; McClelland & Elman, 1986; Norris, 1994) suggest that multiple lexical candidates are activated in parallel and compete for recognition as listeners match incoming auditory information with stored lexical representations. In L2 word recognition, words from both the L2 and L1 may be activated, leading to less efficient recognition when the L1 lacks a corresponding phonemic distinction. This can result in the activation of competing words that are not present in the auditory signal which, in turn, may hinder L2 word recognition.

The lack of group differences in the lexical familiarity test suggests that participants had comparable phonological representations for the target words. This supports the interpretation that group performance differences in the Word Monitoring Task were not due to differences in familiarity with the lexical items but rather related to the processing of orthographic and phonological information.

Another possibility is that these participants were more influenced by the written word, as they showed higher accuracy when the target word contained a <v>. Interestingly, this pattern was only marginal among the Spanish L2 learners but significant among native speakers and Korean L2 leaners. While native speakers demonstrated high accuracy across all conditions, they were particularly accurate when they heard a labiodental sound but saw a word containing the letter <b>. This interpretation aligns with previous studies that have documented the influence of orthography on the pronunciation of [v] (Cartagena, 2002; Stevens, 2000; Takawaki, 2012; Torres Cacoullos & Ferreira, 2000; Trovato, 2017). In conclusion, the L1 phonemic inventory constrains speech discrimination, and the resulting lexical competition, modulated by visual-orthographic information, seems to account (at least in part) for the pattern of results observed in Experiment 2.

3. General discussion and conclusion

This study aimed to investigate how perceptual challenges in distinguishing between phonemes might affect word recognition in L2 learners and how these challenges could be influenced by orthographic cues, particularly focusing on the /b/-/v/ distinction. The findings reveal that both Spanish- and Korean-speaking learners of English do face difficulties distinguishing between [b] and [v], consistent with previous research that attributes this challenge to the allophonic nature of these sounds in Spanish (e.g., Abramson & Lisker, 1970; Werker & Tees, 1984) and its non-existence in Korean (e.g., Cho, 2009). In standard Spanish, [b] and [v] are not distinct phonemes, which complicates their differentiation in English, where they function as separate phonemes. In contrast, the Korean phonemic inventory does not include labial fricatives and consists solely of labial stops: /p/, /ph/, /p'/, and the allophone [b]. In line with current models of L2 phonetic perception, such as the Speech Learning Model (SLM) (Flege, 1995) and the Perceptual Assimilation Model (PAM) (Best, 1995), these findings highlight the impact of L1 phonemic inventories on L2 perception. Both models propose that perceived similarity between L1 and L2 sounds influences the difficulty of acquiring non-native contrasts, and this study's results align with those predictions.

However, this study extends these findings by demonstrating that these perceptual issues can hinder word recognition, particularly when visual-orthographic cues are involved. The word monitoring task revealed differences in how participants processed mismatched phoneme-grapheme pairs. Spanish learners showed a marginally significant effect in the mismatched condition where <b> was seen but /v/ was heard, likely reflecting interference from Spanish orthographic conventions, where <b> and <v> are phonetically indistinguishable (Cartagena, 2002; Stevens, 2000; Takawaki, 2012). In contrast, Korean learners did not exhibit this effect, potentially due to their limited exposure to Roman grapheme-phoneme mappings. The mismatch condition suggests that learners, particularly L2 learners, may have weaker mental representations of the fricative /v/, which is either allophonic or absent in their L1 phonemic inventories. This mismatch between /b/ and /v/ highlights the challenges learners face when processing contrasts involving phonemes not robustly represented in their native languages (Abramson & Lisker, 1970; Flege & Hillenbrand, 1984; Hayes-Harb & Masuda, 2008). While these findings underscore the influence of orthographic conventions on L2 learners' phoneme perception and word recognition, the evidence is limited to specific conditions. Further research is needed to explore whether these effects generalize to other mismatched grapheme-phoneme pairs and to disentangle orthographic influence from phonological and experiential factors.

A methodological consideration in this study is the inclusion of fillers in the AXB task. These fillers were originally incorporated to prevent participants from focusing exclusively on the /b/-/v/ contrast, with the aim of encouraging a more natural engagement with the task. However, as AXB tasks typically do not require fillers to assess perceptual sensitivity, their necessity in this context remains open to discussion. While alternative approaches, such as control items, could have been included to assess task reliability, the methodological choice to use fillers reflects an underexplored area in phonetic contrast research. Future studies could examine whether and how the presence of fillers influences perceptual outcomes in AXB tasks. In the meantime, the results of this study should be interpreted with this consideration in mind.

This study offers insights into how orthographic systems influence L2 phonological perception by comparing Spanish and Korean learners. Spanish participants showed a marginal effect, reflecting their orthographic convention of treating <b> and <v> as interchangeable, while Korean learners, whose L1 lacks both sounds, often substituted them with the lenis stop /p/ (Kong et al., 2011; Lee & Jongman, 2012; Brown & Yeon, 2015). Exposure to Romanization systems may have familiarized Korean learners with Roman letters, creating associations between English phonemes and Romanized counterparts (King, 2002; Sohn, 2001; Park, 2009). This comparison highlights how native and borrowed orthographic systems shape L2 phonological processing and underscores the complex role of orthography in language learning. Interestingly, the proficiency level did not significantly affect the results, which contrasts with some earlier studies (e.g., Takawaki, 2012; Timm, 1976). This might suggest that the influence of L1 phonemic inventory on L2 perception is robust and persists across different proficiency levels. However, this finding should be interpreted with caution, as the proficiency test used primarily measured visual language skills, potentially overlooking differences in auditory perception. Future research should explore whether assessing proficiency through more aurally focused tests might reveal a stronger correlation between proficiency and L2 phonemic contrast recognition and also try to test a broader range of proficiencies.

In conclusion, this study contributes to our understanding of how orthographic systems shape phonological perception and word recognition in L2 learning. By comparing Spanish and Korean learners, it underscores the complex interplay between written and spoken language in L2 acquisition. The findings suggest that orthographic forms can have a significant impact on L2 learners' ability to distinguish and recognize phonemic contrasts, highlighting the importance of considering both auditory and visual modalities in L2 instruction and assessment.

References

Abramson, A. S., & Lisker, L. (1970). Discriminability along the voicing continuum: Cross-language tests. In B. Hála, M. Romportl, & P. Janota (Eds.), *Proceedings of the 6th International Congress of Phonetic Sciences* (pp. 569–573). Academia.

Baayen, H. R. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge University Press. https://doi.org/10.1017/CBO9780511801686

Bassetti, B. (2017). Orthography affects second language speech: Double letters and geminate production in English. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(11), 1835–1842. https://doi.org/10.1037/xlm0000417

Bassetti, B., Hayes-Harb, R., & Escudero, P. (2015). Second language phonology at the interface between acoustic and orthographic input. *Applied Psycholinguistics*, *36*(1), 1–6. https://doi.org/10.1017/s0142716414000393

Bassetti, B., Mairano, P., Masterson, J., & Cerni, T. (2020). Effects of orthographic forms on L2 speech production and phonological awareness, with consideration of speaker level predictors. *Language Learning*, *70*(4), 1218–1256. https://doi.org/10.1111/lang.12423

Bassetti, B., Sokolovic-Perovic, M., Mairano, P., & Cerni, T. (2018). Orthography-induced length contrasts in the second language phonological systems of L2 speakers of English: Evidence from minimal pairs. *Language and Speech*, *61*, 577–597. https://doi.org/10.1177/0023830918780141

Best, C. T. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 171–204). York Press.

Broersma, M., & Cutler, A. (2008). Phantom word activation in L2. *System*, *36*(1), 22–34. https://doi.org/10.1016/j.system.2007.11.003

Broersma, M., & Cutler, A. (2011). Competition dynamics of second-language listening. *Quarterly Journal of Experimental Psychology*, *64*(1), 74–95. https://doi.org/10.1080/17470218.2010.499174

Brown, J. D. (1980). Relative merits of four methods for scoring cloze tests. *The Modern Language Journal*, *64*(3), 311–317.

Brown, L., & Yeon, J. (Eds.). (2015). *The handbook of Korean linguistics.* John Wiley & Sons.

Caramazza, A., Yeni-Komshian, G. H., Zurif, E. B., & Carbone, E. (1973). The acquisition of a new phonological contrast: The case of stop consonants in French-English bilinguals. *The Journal of the Acoustical Society of America*, *54*(2), 421–428. https://doi.org/10.1121/1.1913594

Cartagena, N. (2002). *Apuntes para la historia del español en Chile*. Academia Chilena de la Lengua.

Cho, M. H. (2009). Confusion in the perception of English labial consonants by Korean learners. *한국콘텐츠학회논문지 [Journal of the Korea Contents Association]*, *9*(1), 455–464.

Cutler, A. (2015). Representation of second language phonology. *Applied Psycholinguistics*, *36*(1), 115–128. https://doi.org/10.1017/S0142716414000459

Erdener, V. D., & Burnham, D. K. (2005). The role of audiovisual speech and orthographic information in nonnative speech production. *Language Learning*, *55*(2), 191–228. https://doi.org/10.1111/j.0023-8333.2005.00303.x

Escudero, P. (2007). Multilingual sound perception and word recognition. *Stem-, Spraak- en Taalpathologie*, *15*(2), 93–103.

Escudero, P., Hayes-Harb, R., & Mitterer, H. (2008). Novel second-language words and asymmetric lexical access. *Journal of Phonetics*, *36*(2), 345–360. https://doi.org/10.1016/j.wocn.2007.11.002

Escudero, P., & Wanrooij, K. (2010). The effect of L1 orthography on non-native vowel perception. *Language and Speech*, *53*(3), 343–365. https://doi.org/10.1177/0023830910371447

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233–277). York Press.

Flege, J. E., & Hillenbrand, J. (1984). Limits on phonetic accuracy in foreign language speech production. *Journal of the Acoustical Society of America*, *76*(3), 708–721. https://doi.org/10.1121/1.391257

Goto, H. (1971). Auditory perception by normal Japanese adults of the sounds “l” and “r”. *Neuropsychologia*, *9*, 317–323. https://doi.org/10.1016/0028-3932(71)90027-3

Hayes-Harb, R., & Masuda, K. (2008). Development of the ability to lexically encode novel second language phonemic contrasts. *Second Language Research*, *24*(1), 5–33. https://doi.org/10.1177/0267658307082980

Hayes-Harb, R., Nicol, J., & Barker, J. (2010). Learning the phonological forms of new words: Effects of orthographic and auditory input. *Language and Speech*, *53*(3), 367–381. https://doi.org/10.1177/0023830910371460

Hothorn, T., & Everitt, B. S. (2014). *A handbook of statistical analyses using R.* Chapman and Hall/CRC. https://doi.org/10.1201/9781420010657

Hualde, J. (2014). *Los sonidos del español: Spanish language edition.* Cambridge University Press. https://doi.org/10.1017/CBO9780511719943

Jongman, A., Wayland, R., & Wong, S. (2000). Acoustic characteristics of English fricatives. *The Journal of the Acoustical Society of America*, *108*(3), 1252–1263. https://doi.org/10.1121/1.1288413

Kilborn, K., & Moss, H. (1996). Word monitoring. *Language and Cognitive Processes*, *11*(6), 689–694. https://doi.org/10.1080/016909696387105

King, R. (2002). The Korean Language by Ho-Min Sohn (review). *Academia Koreana*, *5*(2), 99–127.

Kong, E. J., Beckman, M. E., & Edwards, J. (2011). Why are Korean tense stops acquired so early?: The role of acoustic properties. *Journal of Phonetics*, *39*(2), 196–211. https://doi.org/10.1016/j.wocn.2011.02.002

Kuhl, P. K. (1991). Human adults and human infants show a “perceptual magnet effect” for the prototypes of speech categories, monkeys do not. *Perception & Psychophysics*, *50*(2), 93–107. https://doi.org/10.3758/BF03212211

Lee, H., & Jongman, A. (2012). Effects of tone on the three-way laryngeal distinction in Korean: An acoustic and aerodynamic comparison of Seoul and South Kyungsang dialects. *Journal of the International Phonetic Association*, *42*(2), 145–169. https://doi.org/10.1017/S0025100312000035

Lively, S. E., Pisoni, D. B., Yamada, R. A., Tohkura, Y. I., & Yamada, T. (1994). Training Japanese listeners to identify English /r/ and /l/. III. Long-term retention of new phonetic categories. *The Journal of the Acoustical Society of America*, *96*(4), 2076–2087. https://doi.org/10.1121/1.410149

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, *25*(1–2), 71–102. https://doi.org/10.1016/0010-0277(87)90005-9

Martínez-García, M. 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

Mathieu, L. (2016). The influence of foreign scripts on the acquisition of a second language phonological contrast. *Second Language Research*, *32*(2), 145–170. https://doi.org/10.1177/0267658315601882

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*(1), 1–86. https://doi.org/10.1016/0010-0285(86)90015-0

Melnik, G. A., & Peperkamp, S. (2021). High-variability phonetic training enhances second language lexical processing: Evidence from online training of French learners of English. *Bilingualism: Language and Cognition*, *24*(3), 497–506. https://doi.org/10.1017/S1366728920000644

Mok, P. P. K., Lee, A., Li, J. J., & Xu, R. B. (2018). Orthographic effects on the perception and production of L2 mandarin tones. *Speech*

Communication, *101*, 1–10. https://doi.org/10.1016/j.specom.2018.05.002

Norris, D. (1994). Shortlist: a connectionist model of continuous speech recognition. *Cognition*, *52*(3), 189–234. https://doi.org/10.1016/0010-0277(94)90043-4

Pallier, C., Colomé, A., & Sebastián-Gallés, N. (2001). The influence of native-language phonology on lexical access: exemplar-based versus abstract lexical entries. *Psychological Science*, *12*(6), 445–449. https://doi.org/10.1111/1467-9280.00383

Park, J. S. Y. (2009). *The local construction of a global language: Ideologies of English in South Korea.* Mouton de Gruyter. https://doi.org/10.1515/9783110214079

Polka, L. (1992). Characterizing the influence of native experience on adult speech perception. *Perception and Psychophysics*, *52*, 37–52. https://doi.org/10.3758/bf03206758

Polka, L., & Werker, J. F. (1994). Developmental changes in perception of nonnative vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 421–435. https://doi.org/10.1037//0096-1523.20.2.421

R Core Team (2009). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing. https://www.r-project.org/

Rafat, Y., & Stevenson, R. A. (2019). Auditory-orthographic integration at the onset of L2 speech acquisition. *Language and Speech*, *62*(3), 427–451. https://doi.org/10.1177/0023830918777537

Shin, J., Kiaer, J., & Cha, J. (2012). The sounds of Korean. Cambridge University Press. https://doi.org/10.1017/CBO9781139342858

Sohn, H. M. (2001). *The Korean language.* Cambridge University Press.

Stevens, J. (2000). On the labiodental pronunciation of Spanish /b/ among teachers of Spanish as a second language. *Hispania*, *83*(1), 139–149. https://doi.org/10.2307/346152

Strange, W., Akahane-Yamada, R., Kubo, R., Trent, S. A., & Nishi, K. (2001). Effects of consonantal context on perceptual assimilation of American English vowels by Japanese listeners. *Journal of the Acoustical Society of America*, *109*, 1691–1704. https://doi.org/10.1121/1.1353594

Tagliaferri, B. (2005). *Paradigm* [Computer software]. Perception Research Systems, Inc.

Takawaki, S. L. (2012). *Orthographic loyalty in the Spanish of northern Mexican speakers* [Master's thesis, Arizona State University]. KEEP. https://keep.lib.asu.edu/items/151287

Timm, L. A. (1976). Three consonants in Chicano Spanish: /x/, /b/ and /d/. *Bilingual Review / La Revista Bilingüe*, *3*(2), 153–162. https://www.jstor.org/stable/25743678

Torres Cacoullos, R., & Ferreira, F. (2000). Lexical frequency and voiced labiodental-bilabial variation in new Mexican Spanish. *Southwest Journal of Linguistics*, *19*(2), 1–17.

Tranter, N. (2000). The phonology of English loan-words in Korean. *Word*, *51*(3), 377–404. https://doi.org/10.1080/00437956.2000.11432504

Trovato, A. (2017). *A sociophonetic analysis of contact Spanish in the United States: Labiodentalization and labial consonant variation* [Doctoral dissertation, University of Texas]. UT Electronic Thesis and Dissertations. http://hdl.handle.net/2152/63721

Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language*, *50*(1), 1–25. https://doi.org/10.1016/s0749-596x(03)00105-0

Werker, J. F., & Lalonde, C. E. (1988). Cross-language speech perception: Initial capabilities and developmental change. *Developmental Psychology*, *24*, 672–683. https://doi.org/10.1037/0012-1649.24.5.672

Werker, J. F., & Logan, J. S. (1985). Cross-language evidence for three factors in speech perception. *Perception & Psychophysics*, *37*, 35–44. https://doi.org/10.3758/bf03207136

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. Infant *Behaviour and Development*, *7*, 49–63. https://doi.org/10.1016/s0163-6383(02)00093-0

Appendix A: AXB Materials

stive	stibe
bædɚ	vædɚ
mɔrbəl	mɔrvəl
gəb	gəv
fɔɪbɚ	fɔɪvɚ
lɛb	lɛv

bænt	vænt
bɛmɪʃ	vɛmɪʃ
gɔbl	gɔvl
bɔɪnt	vɔɪnt
lɔbəlz	lɔvəlz
raɪbIk	raɪvIk

bɛln	vɛln
bæstən	væstən
bæbr	bævr
mɪrbl	mɪrvl
bɛs	vɛs
mibrəz	mivrəz

Appendix B: Word Monitoring Task Materials

Forms with /b/	Forms with /v/
bail	veil
bailed	veiled
bailing	veiling
bails	veils
balance	valance
bale	vale
ballet	valet
bane	vain
banish	vanish
banished	vanished
banishes	vanishes
banishing	vanishing
bat	vat
bats	vats
bend	vend
bended	vended

Forms with /b/	Forms with /v/
bending	vending
bends	vends
best	vest
bet	vet
biking	viking
boat	vote
boats	votes
bolt	volt
bow	vow
bowel	vowel
bury	very
bye	vie
curb	curve
curbed	curved
curbing	curving
curbs	curves

Forms with /b/	Forms with /v/
dribble	drivel
dub	dove
dubs	doves
fibre	fiver
fibres	fivers
gabble	gavel
lobes	loaves
marble	marvel
marbled	marvelled
marbles	marvels
rebel	revel
rebels	revels
robe	rove
robed	roved
robes	roves
robing	roving