


ORIGINAL ARTICLE

# Spanish speakers' English schwar production: Does orthography play a role?

Christine Shea 

Departments of Spanish and Portuguese and Linguistics, University of Iowa, Iowa City, IA, USA  
Corresponding author. Email: [christine-shea@uiowa.edu](mailto:christine-shea@uiowa.edu)

(Received 15 May 2020; revised 8 May 2021; accepted 16 May 2021)

## Abstract

This study examines how input mode – whether written or auditory – interacts with orthography in the production of North American English (NAE) schwar (/ɜ-/ found in *fur*, *heard*, *bird*) by native Spanish speakers. Greater orthographic interference was predicted for written input, given the obligatory activation of orthographic representations in the execution of the task. Participants were L1 Mexican Spanish/L2 English speakers (L2,  $n = 15$ ) and NAE ( $n = 15$ , rhotic dialect speakers). The target items were 10 schwar words and 10 words matched in graphemes to the onset and nucleus of the schwar words (e.g., *bird* was matched with *big*), for a total of 20 items. The degree of overlap between schwar productions across group and input mode (L2 only) was analyzed, followed by a generalized additive mixed model analysis of F3, one of the acoustic cues to rhotacization. Results showed that L2 schwar productions were different from the NAE productions in both the overlap and F3 measures, and the written input mode showed greater L1 orthographic interference than the auditory input mode, supporting the hypothesis that L1 orthography–phonology correspondences affect L2 productions of English schwar words.

**Keywords:** plain vowels; neutral vowels; English schwar; L2 English; L1 Spanish; pronunciation

## Orthography–phonology correspondences in L1 and L2

Research on the interface between orthography and phonology in native speakers has consistently shown that orthographic information is activated in spoken word recognition (Pattamadilok et al., 2014; Qu & Damian, 2017; Seidenberg & Tanenhaus, 1979). Two general accounts of this process exist in the literature. One assumes online cross-activation of orthographic information when phonology is accessed (Cutler & Davis, 2012) and the second assumes the integration of phonology and orthography offline, as a consequence of literacy acquisition (Perre et al., 2009). Early evidence for the online account comes from a study by Roelofs (2006), who investigated how spelling inconsistency in Dutch affected native speakers' spoken word production across oral reading, object naming, and prompt-response word generation. Spelling consistency effects were observed only in the reading task,

© The Author(s), 2021. Published by Cambridge University Press

which, it can be argued, is the task where spelling is the most relevant for successful completion. These results suggest that orthographic effects do not automatically arise during lexical recognition and instead, that listeners cross-activate orthographic codes online, via bidirectional links between orthography and phonology (Qu & Damian, 2019). The offline, or “restructuring” account, on the other hand, assumes that representations are altered during literacy development, resulting in full integration between the two (Montant et al., 2011).

Other researchers have taken a different approach and examined how orthographic effects emerge over the course of learning. In these studies, learners are trained on novel words and objects to see how the presence of orthography affects acquisition and retention of the associated phonology–orthography pairs. Using this methodology, Bürki et al. (2012) examined the acquisition of two sets of novel French words. The target words contained consonant clusters that can be produced with or without a schwa ([pələd] *peloude* vs. [plun] *ploune*). In spoken speech, the reduced form is more common. Participants were trained to associate the spoken reduced forms of the novel words with novel objects and subsequently, they were exposed to the written form of the word once. Importantly, half of the orthographic representations had a schwa while the other half did not. In a naming task, participants produced more schwa segments for items spelled with a schwa and were also slower on producing the words with the schwa variant and in a recognition task, participants were more likely to say that the schwa word was part of the training set when the spelling was consistent with it. According to Bürki et al., these results demonstrate that a single exposure to spelling can change the way words are processed and stored in perception and production. This finding supports an offline link between orthography and phonology and suggests that brief exposure to orthography can serve to restructure phonological representations offline.

In the context of L2 orthography–phonology effects, research has shown that, depending upon the languages and sounds involved, L2 orthographic transfer effects on production and perception can be neutral (Escudero & Wanrooij, 2010), facilitative (Escudero et al., 2008; Showalter & Hayes-Harb, 2013), or negative. Evidence for a facilitatory effect comes from Escudero et al. (2008). The authors trained L1 Dutch/L2 English learners to associate novel English words with novel shapes. The words contained either /æ/ or /ɛ/, a contrast that is very challenging for Dutch learners. Participants were trained on auditory input alone or on auditory + orthographic input that had different graphemes for the /æ/ or /ɛ/ distinction. The results showed that the auditory input group represented the words with the same vowel, while the auditory + orthography input group represented the words with two different vowels, showing that seeing orthographic word forms in a second language influenced the phonological representations of these words. For L2 production, Solier et al. (2019) found effects for training modality (oral vs. written) on the production of five word-final French vowels by L1 Moroccan Arabic learners. Posttest pronunciation was more accurate for participants who had received orthographic input in the training session than for those who had received only auditory input.

As an example of negative orthographic effects, Rafat (2016) points to the problem that English-speaking learners of Spanish have with incorrectly transferring the phoneme /z/, which corresponds to <z> in English, to their L2, in which /z/ does

not exist. In Spanish, [z] exists only as an allophone of /s/, as in words such as <mismo> “same” [mizmo] and <asno> “donkey” [azno]. For English speakers, /z/ is connected to the grapheme <z> while in Spanish, <z> is connected to the grapheme <s> and to the phoneme /s/ and only allophonically (due to voicing assimilation) to [z]. Thus, English speakers tend to produce words like <zapatos> as [zapatos], while the native-speaker target is [sapatos]. While there are a few notable exceptions, such as the <s>, <c>, and <z> graphemes, which correspond to [s] in Latin American Spanish, and the <v> and <b> graphemes which can correspond to variants of [b] and [β], overall Spanish is an orthographically transparent language while English has an opaque phoneme-to-grapheme correspondence system.<sup>1</sup>

Another example of negative transfer comes from a study by Bassetti (2017) that examined L1 Italian/L2 English learners’ production of English consonants represented by double graphemes, such as <rr>, in *berry*, or <nn>, in *inning*. In Italian, geminate consonants are contrastive ([fato] *fato* “fate” vs. [fatto] *fatto* “fact”) and represented by a doubling of the grapheme. In English, double graphemes occur, but do not correspond to phonological geminates. Bassetti’s Italian speakers lengthened the consonants in English words with double graphemes, providing further support for L1 orthography–phonology correspondences influencing L2 production.

Together, these studies show that L2 perception and pronunciation development is intricately intertwined with L1–L2 orthography–phonology correspondences. Such correspondences underlie the way in which L2 learners encode the sounds of their target language and can play a determining role in L2 phonological development (Young-Scholten & Langer, 2015). At least part of the reason for these effects can be attributed to both the timing and context of L2 acquisition. Adult L2 classroom learners are literate in their first language and when acquiring their second language, they are typically exposed to the written and spoken forms of L2 words concurrently, right from the initial class sessions (Rafat, 2016), well before the L2 phonological representations are in place. Thus, L2 learners often encode L2 orthography–phonology correspondences through the perceptual categories of their L1. This may lead to lasting L2 sound–spelling interference in perception and production.

While abundant evidence suggests that orthography serves to restructure phonological representations (Escudero et al, 2008), it is also true that in languages with opaque orthography, such as English, online effects can arise when inconsistencies emerge in the phonological realization of words that overlap in orthographic representations (e.g., *pint-mint*) (Roelofs, 2006; Ziegler and Ferrand, 1998). In languages such as Spanish, *pint-mint* homographs do not occur. This suggests that the emergence of online orthographic effects in word recognition may depend upon the language’s orthographic system, further supporting the language-dependent nature of orthographic effects in lexical encoding.

In the present study, we examine how native speakers of Mexican Spanish produce North American English (NAE) *schwar*, found in words such as *her* [hɜː], *turn* [tɜːn], and *first* [fɜːst]. *Schwar* can be represented with five different digraphs (<her>, <earned>, <shirt>, <work>, <hurt>). This can represent a challenge for Spanish speakers accustomed to transparent orthography–phonology correspondences. In Spanish, vowels are produced consistently more or less across all

contexts (reduced syllables do not occur). For example, the vowel /i/ is consistently produced as a high-front vowel, regardless of its position relative to stress. Moreover, each vowel grapheme represents one sound and when two vowel graphemes occur in the same syllable, they represent a diphthong. For example, Spanish words such as *real* “real” and *pieza* “piece” are pronounced as raising diphthongs. Their cognates in American English *real* and *piece* are pronounced with the monophthongal /i/. When a sequence of two vowel graphemes occurs in different syllables in the same word, the result is hiatus in both Spanish and American English (Spanish: *teatro* [te.a.tro], *theater* [θi.a.tə]). In the particular case of the digraph <ea> in American English, words can be produced with hiatus, such as *theatre* or *idea* or as plain vowels, as in the words *eat* [it] or *beat* [bit], as a vowel + rhotic as in *ear* [iɹ]/[iə] or [iʔ], or even as schwa, in words such as *earn* [ɜn]. The existence of these words, written with graphemes that exist in Spanish and phonologically overlap in some cases with the Spanish diphthong [ea] or hiatus [e.a], adds to the challenge for native Spanish speakers when producing schwa digraphs such as <ear>. In the next section, we present a description of NAE schwa, followed by a brief discussion of rhotics in Spanish.

### **Rhotics, rhotacized vowels, and schwa in English**

NAE speakers of rhotic varieties show considerable inter- and intra-speaker articulatory variability in the production of /ɹ/ (Delattre & Freeman, 1968; Mielke et al., 2016). Two main tongue shapes are used, which result in distinct articulatory – but not acoustic – allophones. One involves articulation of the rhotic with a bunched posterior tongue position and with the tongue tip pointed down (allophonically represented as [ɹ]), while the other involves a more retroflex articulation, with the tongue tip pointing up (allophonically represented as [ɹ̠], Hagiwara, 1995; Mielke et al., 2016). Despite this articulatory variability both within and between speakers, /ɹ/ exhibits a strong degree of acoustic stability (Delattre & Freeman, 1968). Phonologically, the rhotic consonant can appear in prevocalic position as a singleton (/ɹ/ as in *red*), as part of a consonant cluster (/fɹ/ as in *free*) or in the syllable coda ([hiɹ]/[hiə], as in *hear*).

In terms of how rhotic consonants affect preceding vowels, three types of vowels (at minimum) are commonly identified: plain vowel + rhotic, schwa, and unstressed neutral vowel + rhotic.<sup>2</sup> The plain vowel + rhotic occurs when a rhotic sound follows a vowel, but the vowel quality is maintained. For example, in words such as *car*, the vowel [a] retains its quality and can be transcribed as [kaɹ]. There is a lack of full consensus as to whether the postvocalic rhotic is a rhotic vowel or a combination of vowel + rhotic consonant, which means it is also transcribed as a rhotic diphthong with a /ə/ offglide, or [kəɹ̠].

The two other rhotic-affected vowels can be distinguished in terms of stress. The stressed /ɜɹ/, or schwa, occurs in words such as *word* [wɜɹd], *hurt* [hɜɹt], and the first syllable of *further* [fɜɹðə]. The unstressed /ə/ occurs in the second syllable of words such as *heater* [hitə]. For both the schwa and the unstressed /ə/, rhotacization is present throughout the vowel articulation and, importantly, is not due to the influence of the rhotic consonant. Thus, they are often phonemically transcribed

with the vowel symbol and the diacritic for rhotacization /-/ attached to them, indicating the temporal overlap between the rhoticity and vowel (Rogers, 2014).

In the present study, the focus is on the stressed /ɜ-/ schwar. Schwar can occur in word-final position (as in *fur*, *her*) and in preconsonantal position (as in *bird*, *heard*, *word*). Distinct from plain vowel + rhotic, the schwar (and its unstressed counterpart) is articulated with complete rhoticity. In a study carried out by Kuecker et al. (2015), the authors showed that the stressed schwar produced by native NAE speakers exhibited rhotacization across 94% of its articulation (measured by the point at which F3 began to fall) while the plain vowel + rhotic tokens showed only 58% and the unstressed schwar 76%. To determine exactly what a “lower F3” might mean, Hagiwara (1995) examined rhotic productions by California speakers of American English. He calculated a neutral third formant value by averaging the F3 formant values of speakers’ plain (non-rhotic) vowels and comparing this average to the F3 for speakers’ schwar productions. He found that, indeed, the F3 values for the schwar ranged from 56 to 77% of the F3 values for the plain vowels and concluded that rhoticity should in fact be calculated on a scale, based upon a ratio between the F3 for plain vowels and the F3 for schwar, rather than absolute Hz values.

### **Rhotic sounds in Spanish**

Normative Spanish has two rhotic sounds, known as *vibrantes*: /r/, as in *pero* “but,” and /r/, as in *perro* “dog.” While the phonological status of these two sounds has been the subject of many analyses (Bradley, 2019; Harris, 1969; Hualde, 2005), for the present purposes, only /r/ is considered since it is the only rhotic that can occur in coda position in (normative) Spanish. The sound /r/, or tap, occurs in the same phonotactic positions as the American English rhotic, specifically, in word-medial position between two vowels (as in *toro* [t̟oro] “bull”), as the second member of a consonant cluster (as in the word *triste* [triste] “sad”), in syllable coda position word-internally (as in *fuerte* [fuerte] “strong”) and word-finally (as in *correr* [korer] “to run”). The articulation of the tap in Spanish involves a quick single contact with the alveolar ridge and while there is a great deal of variability in its realization across speakers, within speakers, and across dialects (Willis & Bradley, 2008), /r/ is not articulatorily similar to the English rhotic allophones [ɹ] or [ɹ̥]. There is not much (if any) evidence suggesting that L1 Spanish speakers cannot perceptually encode American English plain vowel + rhotic sequences, the schwar, or unstressed vowel + rhotic sounds. That is, they do not have difficulties perceiving the sound in words. Rather, the challenge lies in articulating it.

### **Current study**

As mentioned, schwar exhibits rhoticity throughout its articulation and can be represented by at least five different orthographic symbols. For L1 Spanish speakers whose native language has relatively transparent orthography–phonology correspondences, it is possible that these multiple orthographic representations of the single schwar vowel in NAE could lead to orthographic interference from Spanish. According to online models, such interference should be greater when

**Table 1.** Profile of L2 participants

English experience	Mean (SD)
Age (years)	20.7 (1.4)
Years studying English (years)	7.5 (.7)
Age at which began to study English (years)	13.2 (2.4)
Number of hours of English spoken outside of English classes	3.23 (1.18)
Number of hours spent watching English media (without subtitles)	6.3 (4.1)

the input is in the written form than when it is auditory. Specifically, when the input is written, learners see the graphemes in their L2 that correspond directly to graphemes in the L1, and the L1 phonological categories are activated (Solier et al., 2019). If this is the case, L1 Spanish speakers will produce schwar as plain vowel + rhotic sequences, resulting in recognizable vowel quality that corresponds to the orthography–phonology correspondence in Spanish (or possibly English) in the early phases of the schwar production. When the input is auditory, on the other hand, L1 the orthography–phonology correspondences may be easier to inhibit.

To test this hypothesis, we analyzed the productions of 15 native Mexican Spanish speakers' (L2 speakers) and compared them to the productions of 15 NAE speakers. The schwar vowel is characterized by steady F1 and F2 formant values (reflecting the front, mid-vowel quality) and low F3 formant values (indicating rhoticity) throughout its articulation. The plain vowel + rhotic productions, on the other hand, are characterized by two phases (Hagiwara, 1995): a first phase where the vowel is articulated and a second phase, where rhoticity begins. In the first phase, F1 and F2 formant structure corresponding to the plain vowel can be identified while in the second phase, a fall in F3 indicates the onset of rhoticity.

The study's goals presented two challenges for stimuli design. First, it was necessary to match target schwar words with non-schwar words that had the same vowel graphemes and, importantly, corresponded to a vowel sound in Spanish. Thus, schwar and onset-nucleus matched words were selected to fulfill these criteria (see Table 2). Second, it was impossible to predict whether the L1 Spanish speakers would produce the items with Spanish orthography–phonology correspondences or English correspondences. For example, the vowel in the word *big* is the English vowel /i/, a vowel not found in the Spanish inventory. When producing /i/, Spanish speakers tend to substitute the vowel /i/ (Morrison, 2009), which corresponds to the grapheme <i> in that language. When producing the schwar in the word *bird* [bɜːd], native Spanish speakers could produce the high-front vowel /i/, resulting in [biɪd] or even [biːɪd], which would be closer acoustically to the English word *beard*. This complicates the possible phonetic realizations of the schwar and non-schwar words considerably. To address this, we present two analyses. First, we examined F1 and F2 values across the schwar and non-schwar items to determine whether the participant produced a plain vowel. Second, we examined F3 vowels for the schwar targets to determine the amount of rhoticity present across the vowel. If the participant produces the target word *first* [fɜːst] with a plain vowel, as in

Table 2. Stimuli

Schwar targets			Plain vowel targets		
Orthographic representation	Word	IPA	Orthographic representation	Word	IPA
<er>	her	[hɜː]	<e>	help	[hɛlp]
	were	[wɜː]		wet	[wɛt]
<ear>	earned	[ɜːnd]	<ea>	eat	[it]
	heard	[hɜːd]		heat	[hit]
<ir>	bird	[bɜːd]	<o>	big	[bɪg]
	shirt	[ʃɜːt]		ship	[ʃɪp]
<or>	work	[wɜːk]	<o>	<u>w</u> oman	[wʊmən]
	worth	[wɜːθ]		woke	[wok]
<ur>	hurt	[hɜːt]	<o>	<u>h</u> uman	[hjuːmən]
	turn	[tɜːn]		tube	[tub]

[fɪst] or [fɪːst], there will be measurable formant structure corresponding to the [i] vowel before the rhoticization begins (first analysis) and if the participant produces full rhoticization across the entire target, there will be no fall in F3 values.

To facilitate this, multivariate analysis of variance (MANOVA) was used (F1 and F2 as dependent variables), as it allows us to test hypotheses regarding the effect of one or more independent variables on two or more dependent variables. The output from MANOVA includes a number of summary statistical tests, one of which is Pillai's Trace, or Pillai score (Hay et al., 2006; Kelley & Tucker, 2020; Nycz & Hall-Lew, 2013). The Pillai score describes the separability of two distributions as well as variation within each distribution. Scores close to 0 correspond to overlapping or merged categories and scores close to 1 correspond to distinct categories. While we report the significance for each MANOVA model, the Pillai scores in and of themselves were not directly compared for significance but rather compared along a scale of relative degree of overlap between the F1 and F2 values for each target vowel. The second analysis examined the degree of rhoticization in schwa productions by comparing F3 trajectories across groups and input mode (for L2 speakers). As stated, F3 is lower for schwa than for full vowels (Hagiwara, 1995) and, importantly, for NAE speakers, F3 is low throughout the vowel.

Together, these two analyses allow us to evaluate the degree of overlap between the schwa and plain vowel targets (Pillai score) for each group and determine the degree of rhoticity present in the schwa targets for each group (F3 measurement), across the two input modes. In this way, we can test the hypotheses that first, the L2 speakers will produce plain vowel + rhotic sequences instead of the target schwa and second, where L2 speakers are producing plain vowels + rhotic sequences, the plain vowel portion will exhibit greater orthographic interference from Spanish in the written input mode than the repetition input mode.

## Method

### Participants

Participants were 15 native speakers of Spanish from Mexico (11 male) and 15 native speakers of American English (8 male, speakers of the North or Midland variety, both rhotic). The native American English speakers were undergraduate and graduate students enrolled at a US university in the Midwest (avg. age: 24 years,  $SD = 3.2$ ). They were recorded in a sound-proof booth at a university phonetics lab. The Mexican participants (avg. age: 22 years,  $SD = 2.5$ ) were recruited via word-of-mouth from the same undergraduate program at a public university in Mexico City. All native Mexican–Spanish speakers had learned English as an L2 after 12 years of age, had not spent more than three continuous weeks in an English-speaking country, and did not speak English as part of their regular daily activities. None were studying degree programs related to English or English-language teaching. All recordings were carried out in a sound-attenuated room at their university. Table 1 presents the linguistic background information for the L2 participants:

### Stimuli

Ten words with schwar /ɜ-/ were selected and matched with 10 plain stressed vowel control items (see Table 2) that overlapped in the onset and nuclear graphemes (e.g., <turn> and <tube>). Of the 10 schwar targets, all were monosyllabic and eight had #CVC# syllable structure; the remaining two were #CV#. For the plain vowel targets, seven had #CVC#, one had #CVIC#, and two were bi-syllabic #CVCVC#. One of the plain vowel words was a cognate (“tube” *tubo* in Spanish). Other than *human* and *tube*, the items were non-cognates (or at least non-obvious cognates). While the differences in syllable structure were not ideal, it was determined that matching for graphemes and known vocabulary was the priority. Table 2 presents the word list for the schwar targets and the corresponding words with plain vowels.

### Procedure

Recordings done in the USA were carried out in a sound-proof booth, with a Shure SM58 microphone connected to a Marantz PD671 solid state recorder and then transferred to a laptop computer for analysis. Recordings in Mexico were done with a USB Blue Yeti microphone directly onto a PC laptop computer using Audacity. All recordings were done at a sampling rate of 41000 Hz.

Input mode was counterbalanced across participants and word order was pseudo-randomized such that no more than two words from each condition (schwar vs. plain vowels) occurred together. For the reading task, participants read a list of words presented to them on a computer screen. They were asked to read the words at a comfortable pace, using the carrier phrase “The word is \_\_\_”/ “La palabra es \_\_\_”. For the repetition task, the words were produced by a female Midlands American English speaker. The sound files were embedded into a power point slide show (one item per slide) and participants clicked on the speaker icon to hear the word. Headphones were not available for the Mexican



participants but were not deemed necessary since the study took place in a sound-attenuated room. Participants could listen to each word a maximum of two times before repeating it. Participants only repeated the word, without a carrier phrase and could produce the word only once. If they self-corrected, the first production was taken as the experimental token, unless the pronunciation was too far off-target to be considered representative of the target word. In that case, the corrected, second production was analyzed (six occurrences). They were asked to read the word as soon as it appeared and, in the case of auditory input, as soon as it finished playing. Typically, participants produced the target about 1–2 seconds after the stimulus was presented. After finishing the tasks, participants were asked to verify their understanding of the words. All participants indicated 100% familiarity with all items.

### Analysis

Words were segmented from the sound files using the Montreal Forced Aligner tool (McAuliffe et al., 2017, V. 1.01). Subsequently, vowel onsets and offsets were verified using hand-location in Praat (Boersma & Weenink, 2018). Vowel onset was measured from the onset of periodicity following the release burst of the stop or, where the onset was a fricative, where the periodicity began and intensity increased. For words with glide onsets, vowel onset was measured from an increase in intensity. It was necessary to re-check approximately 30% of [w]-initial words by hand for segmentation accuracy. Vowel offset was registered as a drop-off in intensity (i.e., fall in amplitude, Fox & Jacewicz, 2009) and where there was no longer any identifiable vowel periodicity. For the schwar, boundaries were adjusted to exclude the final coda consonant where necessary. Where the coda consonants were liquids (*help*) or nasals (*turn, earned*), the offset was identified as the beginning of the transition into the coda consonant (Amengual, 2016). All files were then processed with a Praat script to extract word duration, F1, F2, and F3 measurements at seven points in the target vowel: 20–80%, in 10% increments. The first 20% and the last 20% were not included so as to avoid coarticulatory effects with the syllable onset and coda, where necessary. The maximum number of formants extracted was five and the maximum formant value was set at 5 kHz. Because both female and male speakers were recruited, all files were normalized where necessary using the *phonr* package in R (v.3.6.1, R Core Team, 2019) with the Bark Difference normalization procedure (v. 1.0-7, McCloy, 2016).<sup>3</sup>

## Results

### MANOVA models and Pillai scores

In this section, we present the results from three MANOVA models. The first addresses the amount of overlap between the schwar and the plain vowels for each group at the 50% vowel point. If the L2 speakers are producing a plain vowel + rhotic instead of a schwar, we expect to find a high degree of overlap between the two productions at the 50% point. If, on the other hand, the L2 speakers are producing a schwar, there will be little overlap between the two vowels for this group. For the NAE speakers, we do not predict any overlap at all. The second

set of models focuses on cross-group overlap in early (20–40%) and later (60–80%) points of the schwar, to see if the formant values change over the course of the vowel for the schwar targets. The third analysis examines the degree of F1 and F2 overlap in L2 schwar productions across the two input modes, reading and repeating.

To run these models, the *manova* function in R was used and the Pillai score was extracted from those results. As stated above, the Pillai score allows an estimation of the degree of overlap between the F1 and F2 values for each token across the different independent variables. Thus, because we are not interested in the overall significance of the effects or interactions beyond *p*-values (overall significance of the model), we do not report the details of the MANOVA models.

### **MANOVA 1: Degree of vowel overlap across word type by group**

One-way MANOVA models were run for each group to examine the degree of overlap between the production of schwar and plain vowels at the 50% vowel articulation point. The normalized F1 and F2 measurements from the 10 spoken and 10 read words from each group were the dependent variables and the predictor was vowel type (schwar vs. plain vowel).

For the L2 speakers, vowel type had a statistically significant effect on F1 and F2 values ( $F(2, 4197) = 537.62, p = .042$ ). The Pillai score was .20, indicating a high degree of overlap between the schwar and plain vowels. Recall that a Pillai score near zero means no variance is accounted for by the predictor variable and the two groups of responses overlap almost completely while a higher Pillai score (up to a maximum value of 1) means the predictor accounts for a greater amount of variance and the two groups of responses overlap little, if at all (Hay et al., 2006). For the NAE speakers, vowel type also had a statistically significant effect on F1 and F2 values ( $F(2, 4197) = 8276.5, p < .001$ ). However, for the NAE speakers, the Pillai score was .79, indicating almost complete separation of the schwar and plain vowels in terms of F1 and F2 values. These results show first, that differences exist across the two groups in the articulation of the schwar and second, that L2 speakers have greater overlap between the schwar and the plain vowels at the 50% point than the NAE speakers.

Figure 1 presents the F1 and F2 Bark values for the schwar and plain vowels for each group. The different vowel targets are portrayed in separate colors, for ease of interpretation. Figure 1 shows that the F1 and F2 values for the L2 speakers exhibit greater dispersion among the schwar tokens and also overlap with the plain vowel productions, particularly in the case of the words with high-front vowel targets. NAE schwar productions, on the other hand, are clustered tightly together in the same region of the vowel space, indicating little variability across productions.

### **MANOVA 2: Degree of vowel overlap at 20 and 80% across groups**

Another MANOVA model was run on only schwar-target words, across groups, at the 20 and 80% vowel intervals. The dependent variables were F1 and F2 and the predictor was group (L2 vs. NAE). Pillai scores were extracted from each model and compared. The prediction is that the L2 group will show greater overlap with the

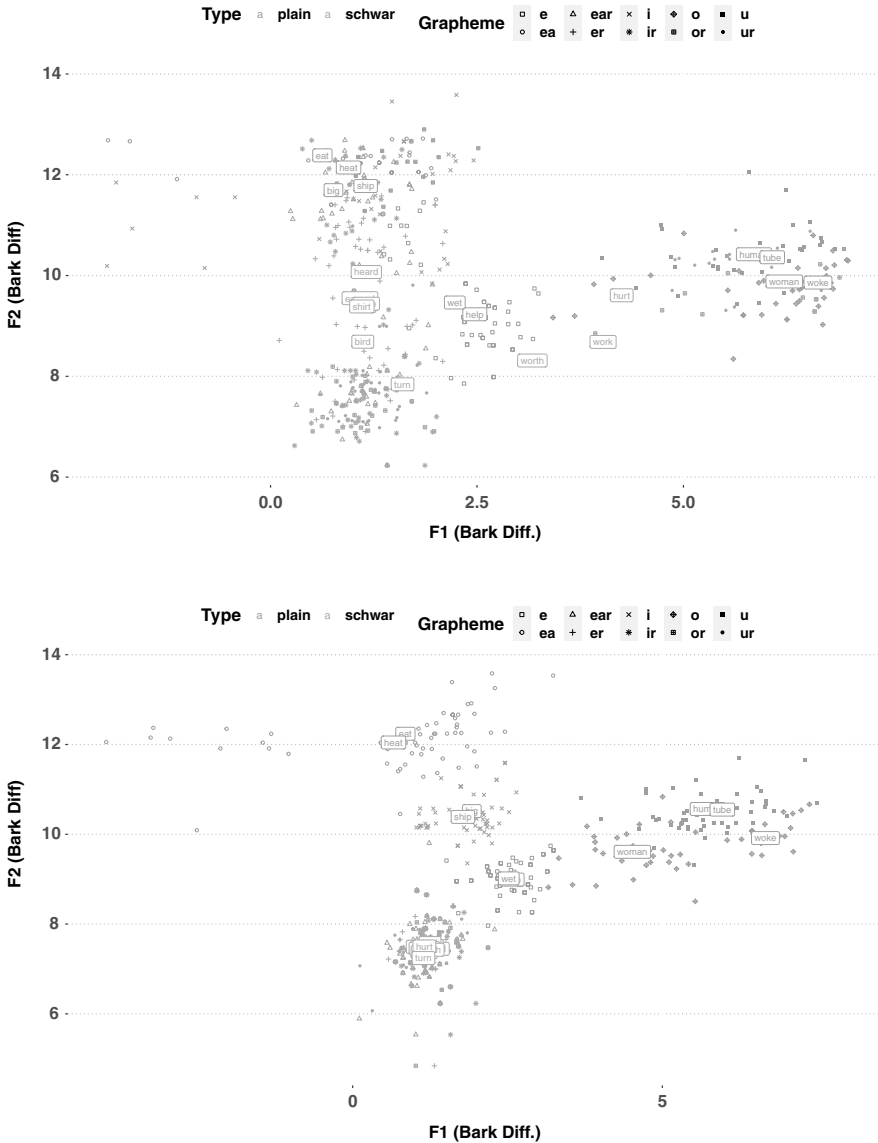


Figure 1. Overlap between schwar and plain vowels for L2 and NAE speakers (50%).

NAE at the 80% vowel point than at the 20% vowel point, at which the L2 productions will be more similar to the plain vowel than the schwar.

The MANOVA revealed significant effects for group on the F1 and F2 values at both the 20% ( $F(2, 4197) = 782.1, p < .001$ ) and 80% vowel point ( $F(2, 4197) = 103.12, p = .03$ ). Figure 2 compares the F1 and F2 values for each group. As can be seen, the L2 productions differ considerably across the two points. At the

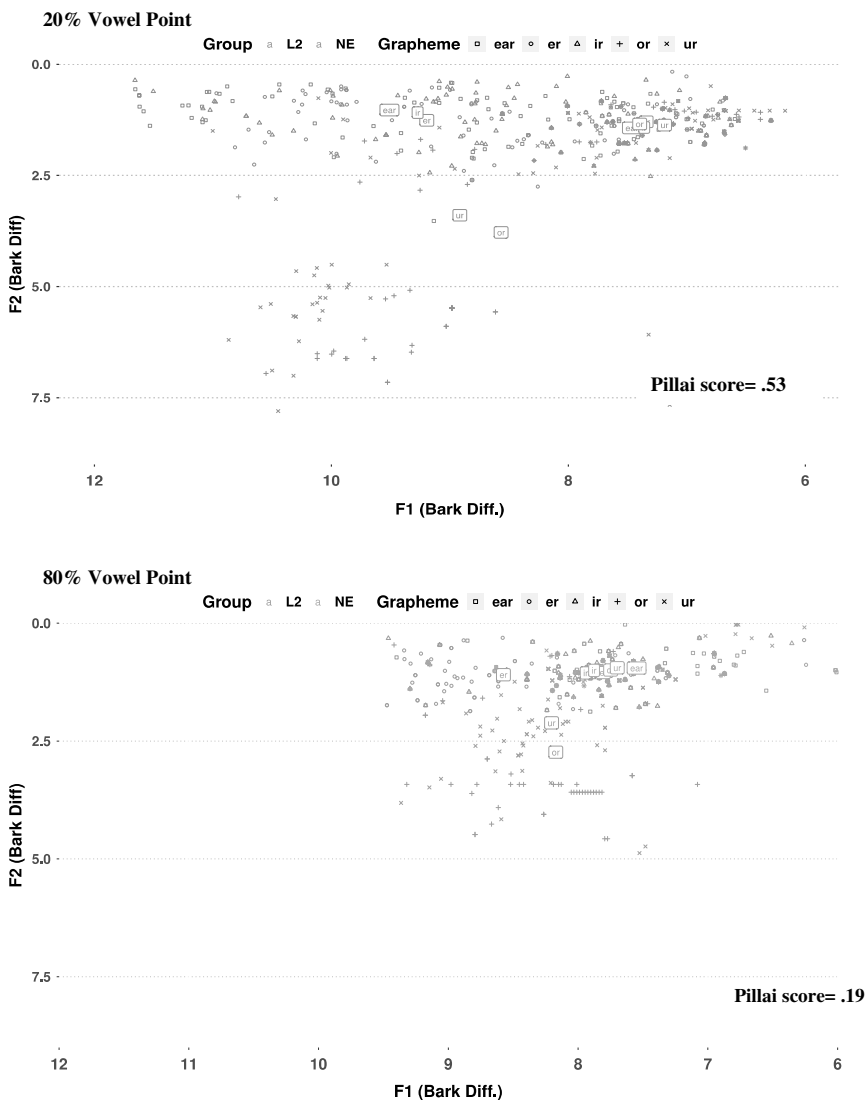


Figure 2. Schwar production across groups at 20 and 80% of vowel articulation.

20% point, there are clusters for each vowel type (high front and back rounded) while this difference diminishes at the 80% point, where the L2 productions are more tightly clustered around the F1 and F2 values of the NAE schwars. This is supported by the Pillai score values. There was less overlap between the productions at the 20% vowel point (Pillai = .53) than at the 80% point (Pillai = .19) between the two groups.

### **MANOVA 3: Degree of L2 vowel overlap between input modes**

The final MANOVA model examined the amount of overlap between input modes for the L2 speakers. The five orthographic representations were separated into two groups, based upon the vowel qualities produced by the L2 speakers. This was determined after inspecting Figures 1 and 2. The first model included schwar targets represented by the orthographic symbols <ear>, <er>, <ir>. The second model included schwar targets represented by <or> and <ur>. The F1 and F2 values are presented for two sets of time points: 20%–30%–40% and 60%–70%–80%. If L2 speakers are indeed producing plain vowel + rhotic sequences, the vowel quality will be observable in the first portions of the vowel and the rhotic will be observable in the latter portion of the vowel. And if input mode affects the degree of vowel quality produced, we predict that the written input will show greater effects for vowel quality earlier in the articulation than the auditory input mode.

In Figure 3, the 20–40% and 60–80% vowel periods are presented for the tokens that correspond to the first group of graphemes, <er>, <ear>, and <ir>. At the early vowel periods, two clusters of tokens can be identified, corresponding to the different input modes. The Pillai score is .48, indicating a moderate degree of separation between the early and late articulation distributions. Clear differences can be observed for each input mode at the 20–40% phase: the tokens produced after the auditory input are clustered together in the vowel space that corresponds more closely to the schwar region while the tokens produced after the written input are clustered more closely in the region of the plain vowel. For the later vowel periods, the degree of overlap is greater (Pillai = .09), indicating less effect for input mode.

Figure 4 presents the same values for the schwar targets represented by the graphemes <ur> and <or>. The overlap between input mode is more pronounced for these tokens at both the early (Pillai = .28) and later periods (Pillai = .02) than for <er>, <ear>, and <ir> graphemes.

### **Discussion of vowel overlap results**

The first MANOVA model confirms that the L2 speakers' productions showed greater overlap between the schwar and plain vowels at the 50% interval of the vowel than the NAE speakers. The second set of models nuanced this finding and showed that the L2 speakers had less overlap with the NAE schwar vowels at the 20% interval of the vowel than at the 80% interval. The third set of models teased apart the distinctions between the different graphemic realizations of schwar and their effect on L2 productions. The items represented orthographically by the sequences <er>, <ear>, and <ir> exhibited F1 and F2 values closer to [i], specifically, higher F2 and lower F1, at the initial portions of the vowel than schwar targets represented by <ur> and <or>. These results confirm that the vowel quality that corresponds to the written symbol in the schwar target affects L2 speakers' productions. This can be clearly seen in Figure 3, where the transfer of the vowel quality to the <ir>, <ear>, and <er> words is clear at the early periods of the vowel articulation. For the <ur> and <or> vowels, on the other hand, there is much greater overlap between the two modes across the early and late periods of the vowel. Both /u/ and /

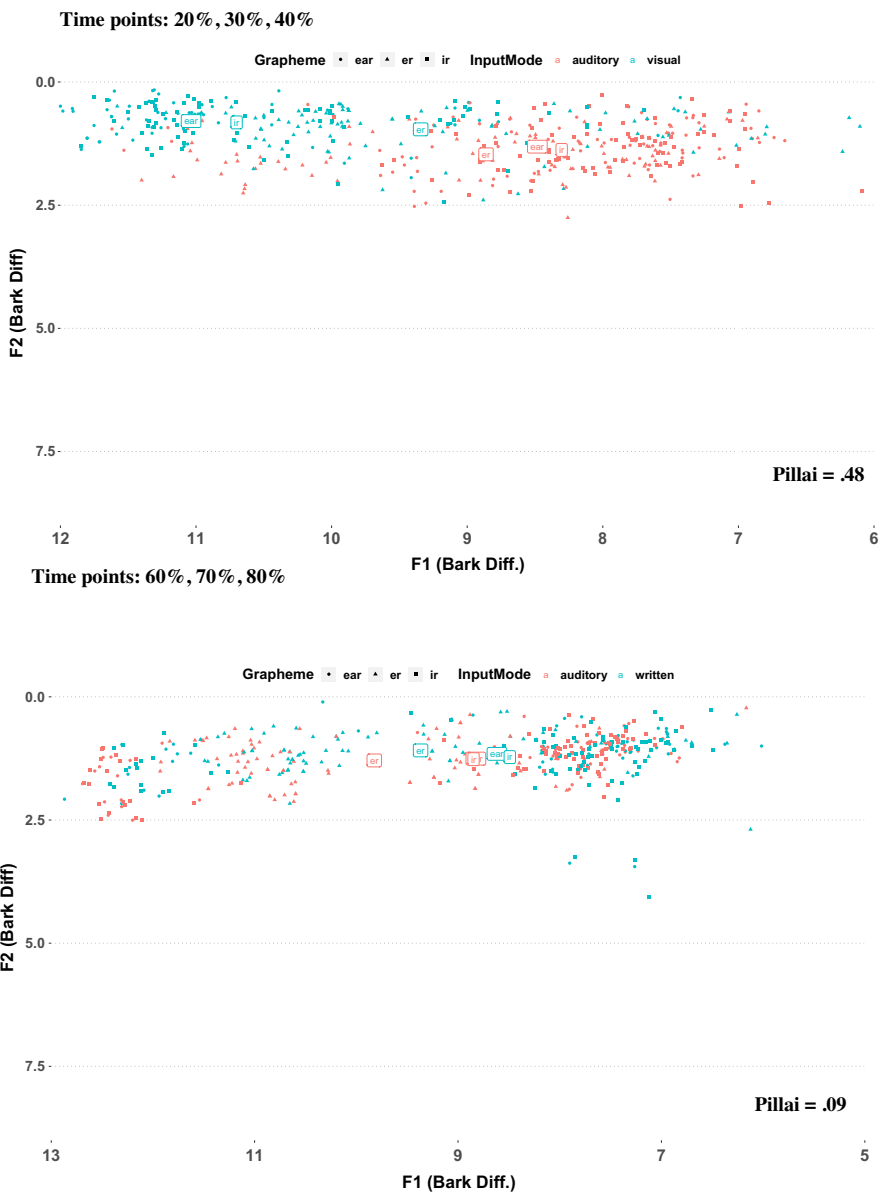


Figure 3. L2 schwar production across modes <ear> <er> <ir>.

o/ are rounded vowels and the rhotic consonant in NAE has been classified as labial under certain phonological analyses (Walker & Proctor, 2019). For both rounded and rhotacized vowels, F3 is lower than for unrounded and plain vowels. The lip-rounding that is involved in the production of rounded vowels lengthens the

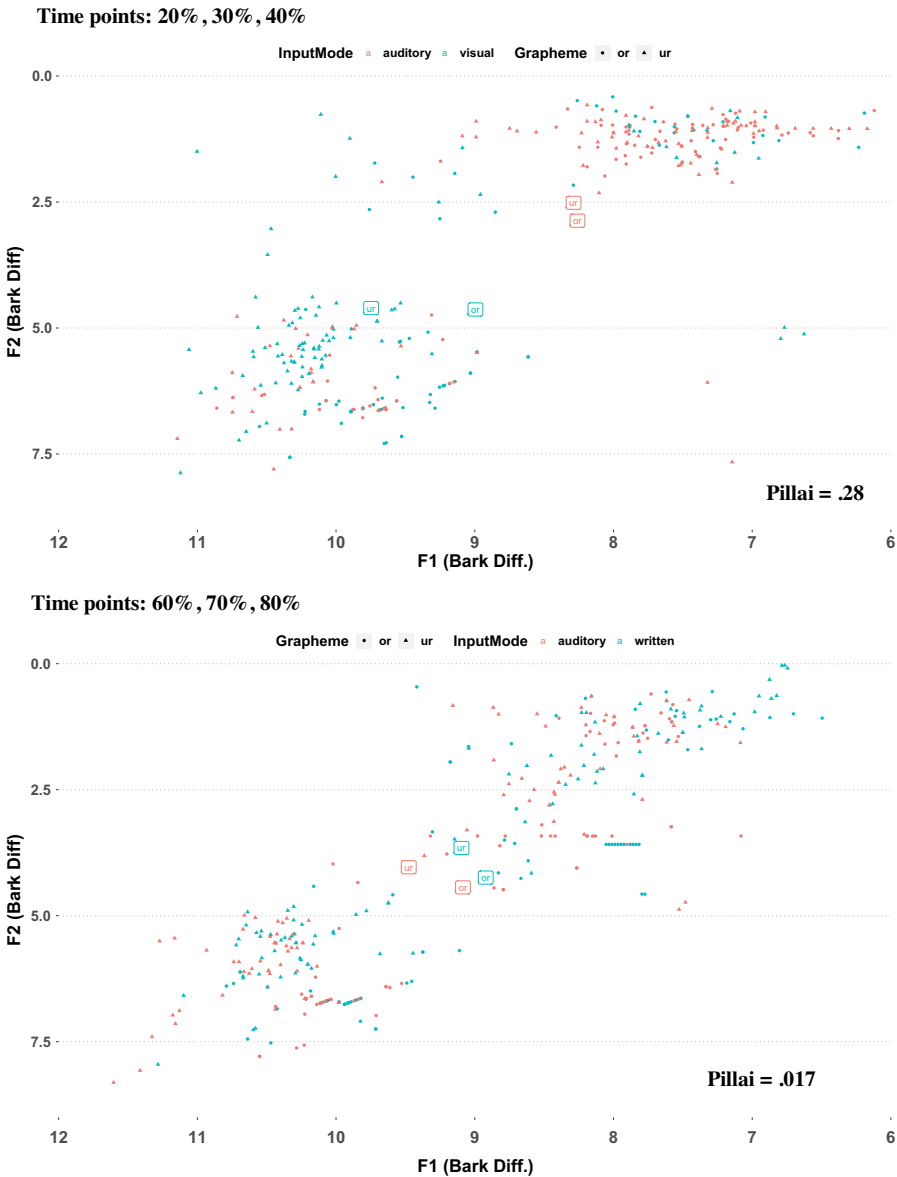


Figure 4. L2 schwar production across modes <or> <ur>.

vocal tract, lowers F2 and F3 overall, and brings their values closer together, which is similar to the overall effect of the rhotic consonant in English.

While Pillai scores are valuable in portraying the degree of vowel overlap, they cannot capture the dynamic nature of vowel articulation across the entire segment or changes in the degree of rhoticity. The comparisons across the first and second periods of schwar targets suggest that the L2 speakers produce F1 and F2 values

closer to those of *schwar* in the later portion of the vowel articulation. In the following section, we examine the F3 trajectory across the L2 and NAE speaker productions. In target-like *schwar* productions, F3 remains low (close to F2) throughout the duration of the vowel (Hagiwara, 1995; Hillenbrand, Getty, Clark & Wheeler, 1995). Given the results from the overlap values presented in the preceding section, F3 is predicted to initiate at a higher point for the L2 speakers and fall over the course of the vowel. For NAE speakers, F3 is predicted to remain low over the entire duration of the vowel.

### **F3 trajectories**

F3 trajectories were analyzed using generalized additive mixed models (GAMM, Sóskuthy, 2017; Winter & Weiling, 2016). Because the F3 trajectory is predicted to be non-linear, traditional linear mixed regression models are not suitable. GAMMs avoid the problems that traditional linear regression models have with capturing non-linear relationships by employing “smooth” terms alongside parametric terms (Sóskuthy, 2017). Another important feature of GAMMs is that they can account for dynamic analyses rather than focusing on a single point of the vowel articulation. Furthermore, GAMMs also allow a comparison between the height and shape of the F3 trajectories across the full articulation of the vowel under different conditions. As the vowel overlap measures presented above reveal, changes in *schwar* production occur across the entire articulation of the target vowel. Two GAMMs are presented in the next section. The first examines differences across groups in terms of the F3 trajectories. The second examines changes in the F3 trajectories across input mode for the L2 speakers only.

#### *Model specification: F3 trajectories across group*

Statistical modeling was carried out using the *bam* function in the *mgcv* package in R (Wood, 2019). The dependent variable for all GAMM models consisted of the non-normalized F3 formant values (Wieling, 2018). The data were filtered to include only the *schwar* vowels for both groups, giving a total of 4,200 values (7 time points  $\times$  10 words  $\times$  2 modes  $\times$  30 participants).

For the GAMM, Group was included as a parametric factor to estimate the constant difference between the L2 and NAE groups on the F3 dependent variable. To examine changes over the seven time points (20–80%), smooths were also included. Smooths allow the incorporation of non-linear patterns in the data, if any exist. They combine a pre-specified number of base functions, called knots, which reflect the number of baseline functions used to model the non-linear patterns. In all the models presented here, each smooth used seven knots for the base functions, since there were seven points at which F3 was measured and a high degree of non-linearity in the data was not expected (Sóskuthy, 2017). A non-linear smooth for Time was included, along with a smooth to model the difference between the two levels of Group across Time. To do so, Group was converted to a binary, ordered variable (Sóskuthy, 2017) and the smooth was specified with “by = Group.ord”. Both smooths used the bs = “cr”, or cubic root specification. Finally, because each group could be producing the items at different rates, another smooth



**Table 3.** Parametric terms of GAMM for Group

Parameter	<i>B</i>	<i>SE</i>
Intercept	1751.7	10.86
Group	-68.6	15.16

with Group.ord and Duration was included. To verify model fit and justify the additional predictors and smooths, the function *compareML* was used (models differing in fixed effects can only be compared using ML estimation) from the *itsadug* package (van Rij et al., 2020). Comparisons revealed that the best model fit included the parametric variable Time and two smooths, one for Time and one for Time by Group. The Group and Duration smooth did not improve the fixed effects fit of the model and was dropped.

For the random effects, random intercepts were included for Speaker, which were allowed to vary across Word (random slope). A second random effects smooth included random intercepts for Speaker and random slopes for Duration. This was included to account for possible effects of duration on F3 shape across speakers. The random intercepts and slopes smooths were run with the basis “re”. Finally, a random factor smooth (basis = “fs”) for Time and Speaker was also included that modeled any non-linear differences that may exist in the data over time, with respect to the time pattern observed for each speaker (Wieling, 2018, p.15). To test the fit of the random effects, AIC values were calculated using *compare* (fREML) and compared across models that included all possible combinations of the random effects. The model was then checked for autocorrelation by means of the *acf\_resid* function in the *itsadug* package. Adjustments were made using the function *bam* (*rho* and *AR.start*) to correct for autocorrelation (*rho* = .678).

### *F3 trajectories across group results*

The results from the first model comparing F3 trajectories across groups for the *schwar* targets are presented in Tables 3 (parametric terms) and 4 (smooth terms). The final model explained 42% of the variance, with an estimated  $R^2$  of .41.

Table 3 presents the parametric terms. In addition to the intercept, the significant coefficient for Group captures the overall difference in the height of the trajectories. The first predictor in Table 4 is the reference smooth, which is fit to the value of L2 speakers. Next, is the difference smooth that captures the differences between the trajectories for the L2 and NAE speakers. The results from these two smooths show that there is a significant difference between the trajectories for the two groups. The edf (effective degrees of freedom) for both was greater than 1, indicating that the difference between the two trajectories was non-linear. Smooths 3, 4, and 5 capture the model’s random effects. Smooth 3 shows that speakers differ significantly in the shape of their F3 trajectories across words. Smooth 4 shows that Duration differs across speakers and affects the shape of the trajectories. Finally, the shape of the trajectories also changes across Time for each Speaker.

Figure 5 provides the plots for the F3 trajectories by Group (5a) and the estimated differences in F3 trajectories by group (5b). As can be seen in Figure 5a, the

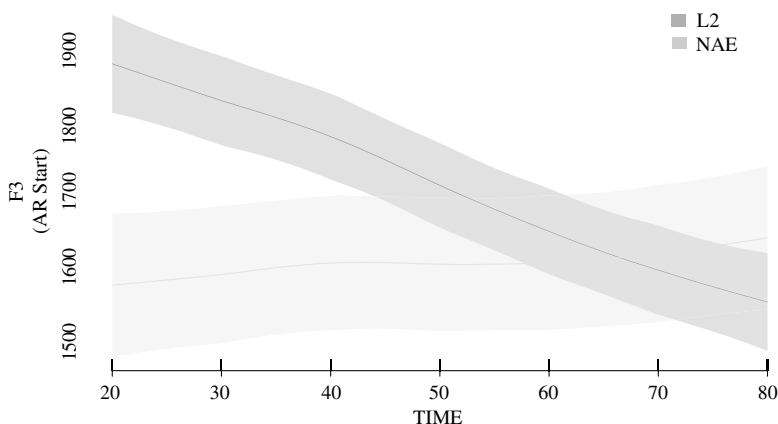
**Table 4.** GAMM smooth terms for Group

Term	edf	Ref.df	F	Sig.
1. Reference smooth: L2	3.3	3.99	90.2	<.001 ***
2. Difference smooth: Group	2.33	2.56	167.8	<.001 ***
3. Random slopes and intercepts: Speaker, Word	20.03	298.00	4.45	<.001 ***
4. Random slopes and intercepts: Speaker, Duration	2.01	30.00	.672	.043 *
5. Factor smooth: Time, Speaker	33.5	206.00	1.756	<.001 ***

Note. The final model fits F3 as a function of Time (20–80% of the vowel).  $R^2 = .41$ ; deviance explained = 42.2%;  $n = 4, 200$ ; edf = effective degrees of freedom; Ref.df = reference degrees of freedom.

\*\*\*  $p < .001$ .

\*  $p < .05$ .

**Figure 5a.** F3 trajectories by Group.

trajectories show a non-linear trend over Time and the L2 speakers show a higher initial F3 value that falls as the articulation of the schwa progresses. The NAE group shows no such trend and also shows a larger confidence interval, reflecting greater inter-speaker variation. The difference plot shows two significant periods of difference, between 20–52% and 77–80% of the vowel.

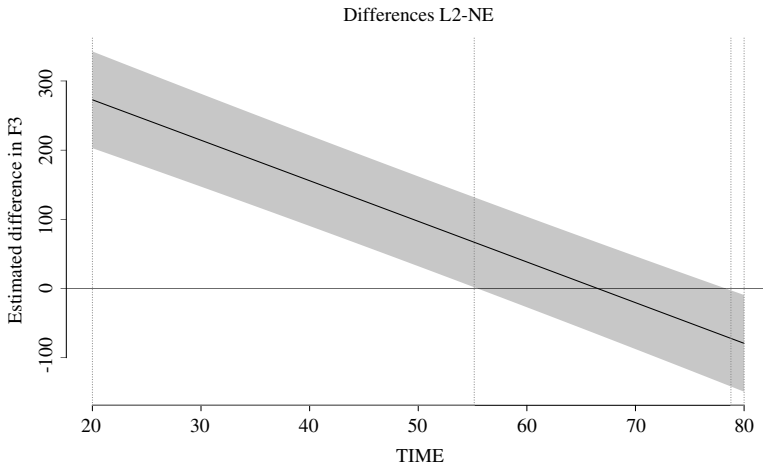
#### *Model specification: F3 trajectories across input mode*

The second GAMM also included F3 as the dependent variable. The data were filtered to include only schwa vowels for the L2 speakers, giving a total of 2,100 values (7 time points  $\times$  10 words  $\times$  2 modes  $\times$  15 participants).

Input mode was included as a parametric factor to estimate the constant differences between reading and repetition on F3. The reference smooth included Time and a difference smooth for Time by Input mode (converted to binary, ordered variable, see above) was also included, to compare the differences between reading and

**Table 5.** Parametric terms of GAMM for L2 mode

Parameter	<i>B</i>	<i>SE</i>
Intercept	1801	21.2
Mode	-128.4	41.2

**Figure 5b.** Differences in F3 trajectories across groups.

repeating across time. Another smooth for Input mode and Duration was also included. Model fit for the fixed factors was carried out using *compareML* and revealed that the best model fit excluded the smooth for Input mode and Duration, which was subsequently dropped. For the random effects structure, a combined Orthography  $\times$  Input mode grouping variable (“OrthMode”) was created (Wieling, 2018, p.15) to capture the effect of Orthography at each level of Input mode, across Time. To test the fit of the random effects, AIC values were calculated using *compare* (fREML) and compared across models that included all possible combinations of the random effects. The model was then checked for autocorrelation by means of the *acf\_resid* function in the *itsadug* package.

### *F3 trajectories across input mode results*

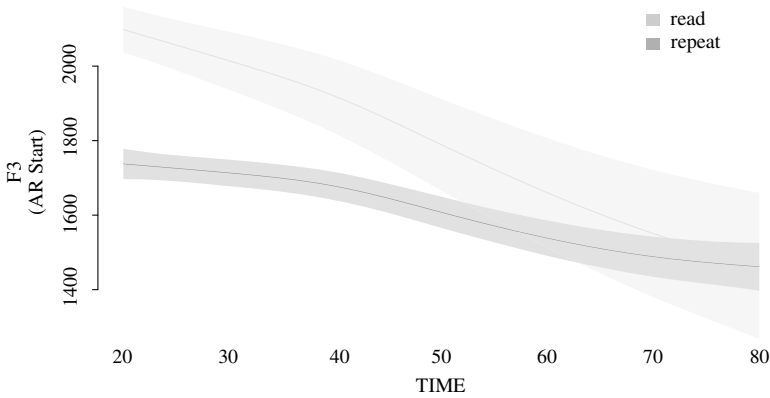
The results from the second model comparing L2 F3 trajectories for *schwar* across input mode are presented in Tables 5 (parametric terms) and 6 (smooth terms). The final model explained 36.5% of the variance, with an estimated  $R^2$  of .35.

The first predictor in Table 6 is the reference smooth, fit to the value of *read*. The second predictor is the difference smooth that captures the difference between the trajectories for the two modes. The differences are significant (this is portrayed in Figure 6b), indicating that the L2 speakers’ productions of the *schwar* differed across input mode. Finally, the third predictor is a factor smooth and captures the effect of Orthography on F3 at both levels of Input mode, across time. The significance of

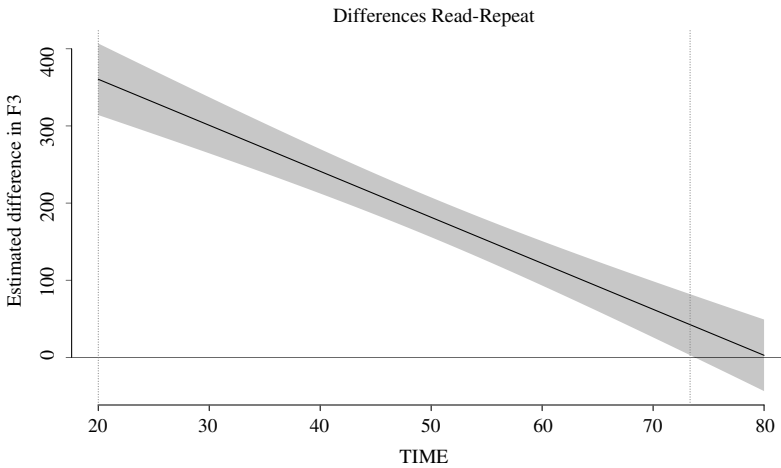
**Table 6.** GAMM smooth terms for L2 mode

Term	edf	Ref.df	F	sig.
1. <i>Reference smooth: read</i>	3.38	4.01	23.2	<.001 ***
2. <i>Difference smooth: Mode</i>	2.98	2.04	19.76	<.001***
3. <i>Random smooth: Time, OrthMode</i>	7.67	8	24.58	<.001***

Note. The final model fits F3 as a function of Time (20–80% of the vowel).  $R^2 = .35$ ; deviance explained = 36.2%;  $n = 2,100$ ; edf = effective degrees of freedom; Ref.df = reference degrees of freedom. \*\*\*  $p < .001$ .



**Figure 6a.** F3 trajectories by Mode (L2).



**Figure 6b.** Differences in F3 trajectories across Mode (L2).

this predictor indicates variability in the effect of Orthography is not constant, which is further supported by the broad confidence interval bands that can be seen in Figure 6.

The trajectories show a non-linear trend over time and a difference between the two modes. Specifically, F3 initiates at a higher value for the read tokens and the right tail shows a broader confidence interval towards the end of the vowel period as well. As can be seen in the difference plot (Figure 6b), the period of significant difference in the F3 trajectories across input mode extends from the beginning of the vowel to over 70% of its duration. These results are consistent with the observations made in the vowel overlap analysis and also with the results from the inter-Group analysis presented in the preceding section.

### **Discussion of GAMM results**

The GAMM allowed us to examine differences in height and shape of the F3 formant across the articulation of the target vowels. Since F3 formants are non-linear, a regular linear mixed model was not appropriate. The results from the GAMM comparing F3 trajectories across Group show significant differences in both height and shape of the trajectories between the L2 and NAE speakers (significant parametric value for Group and significant difference smooth). As shown in Figure 5, the regions of significant difference were predominantly located in the first part of the vowel time periods, consistent with the overlap data presented in the previous section. The L2 speakers show significantly different F3 trajectories for *schwar* words and this difference arises in the first periods of the vowel, during which the F3 formant is significantly higher than that for the NAE speakers. The GAMM also revealed significant differences across Speaker and Word as well as for Speaker and Duration.

The goal of the second GAMM was to compare changes in F3 formant values across time for the L2 speakers with respect to input mode (written vs. auditory). The results revealed significantly higher F3 values in the initial portions of the vowels for the written tokens than for the auditory tokens. The differences across input mode were maintained across Time (significance of the difference smooth). Another important result revealed by the model was the significance of the Time, OrthMode random smooth. Recall that this variable was created by crossing input mode with the different orthographic representations of *schwar* and then examining how this interaction changed across Time. The results reveal that the different representations did affect the F3 trajectories differently for written versus auditory input tokens.

### **General discussion**

The goal of this study was to examine the production of the NAE *schwar*, /ɜ:/, present in words such as *fur*, *her*, *earn*, *turn*, and *word*, by native speakers of Spanish, under two different input mode conditions, written and auditory. L1 Spanish-speaker productions of NAE *schwar* were predicted to exhibit plain vowel quality at the initial stages of the vowel articulation rather than the full rhoticity present in native NAE speaker productions due to negative interference from L1

orthography–phonology correspondences. This effect was expected to be greater in the written input mode than in the auditory input mode. The schwar productions were compared for degree of overlap in F1 and F2 values across groups. The results showed that at the 50% vowel articulation point, there was little overlap between the L2 and NAE speakers but this changed when articulations were compared across the early and later portions of the schwar targets. While at the early stages of vowel articulation, the L2 productions overlapped very little with the NAE's productions, at later stages the overlap increased across all targets. Overall, at the earlier stages, L2 productions were closer to a plain vowel + rhotic. This effect was more pronounced for the written input mode than for the auditory input mode. Input mode had an effect on L2 schwar productions.

The GAMM analysis examined changes in F3 across the articulation of the vowel by group and by input mode for the L2 participants. The schwar produced by the NAE speakers exhibited consistently low F3 values throughout the articulation of the vowel. For the L2 speakers, F3 started higher and fell over the course of the vowel articulation. The GAMM analysis also showed a significant amount of individual variability across the production of the formant values and duration for the schwar targets (significance of the random smooths in model 1) within and across speakers, suggesting that the effect of input mode on F3 trajectories is not uniform across all speakers or across all items. In terms of orthography–phonology correspondences, such individual differences could be due to L2 proficiency (which was not measured here; see Darcy et al., 2015 for individual differences).

In terms of situating these results in the broader literature, there are two main approaches accounting for the locus of orthography–phonology interactions during speech perception and production. Online models claim that orthography is linked to phonology through bidirectional correspondences (Cutler & Davis, 2012), while offline models maintain that orthography is part of phonological representations, and that literacy can restructure representations in the mental lexicon, which, in turn, can restructure perception (Bürki et al., 2012; Taft, 2006). The results from studies such as Escudero et al. (2008) show that orthography can affect phonological encoding when new L2 words are being acquired.

In the present study, we compared across input modes and found that visual input resulted in greater L2 orthographic interference than auditory input. It is possible, of course, that L2 lexical representations are not fully integrated but L1 representations may be (supported by the finding here that there were no input mode effects for the NAE speakers). If this is indeed the case, the task facing L2 learners is one of *inhibiting* L1 connections and the differences across input modes could be due to the relatively greater difficulty in inhibiting L1 orthographic effects when participants are confronted with written input. In the written input mode, explicit orthographic input with direct L1 correspondences activated L1 phonological representations to a greater extent than L2 auditory input because in the latter context, there were no visual stimuli to directly activate the L1 orthography–phonology correspondences. In other words, seeing the graphemes (which have transparent L1 phonological correspondences) makes it more difficult to inhibit the L1 correspondences. Another, related explanation could lie in the timing of orthographic effects (rather than explicit activation vs. non-activation). Seeing the written input could simultaneously activate orthography and phonology, while auditory

input may first activate phonology and subsequently, orthography. For the auditory input, the L1 orthographic codes may be accessed later, and therefore not exert the same degree of influence while for the written input, orthography may be accessed at the same time as phonology, and exert a greater influence upon the subsequent productions.

Thus, while participants' L1 representations could be fully integrated, their L2 representations, we would argue, are not and are instead modulated by input mode. This is perhaps unsurprising given that the L2 learners in this study were late bilinguals who acquired English as an L2 well after becoming literate in Spanish (for Portuguese–English bilinguals, see Roberto Gonçalves & Silveira, 2020). For this particular group, Spanish correspondences are overwhelmingly stronger than those for English.

It is also conceivable that the nature of the phonological representations themselves may play a role in the extent to which L1 orthography affects L2 production. For most L2 learners, the native language is the means through which literacy is taught and L2 literacy development occurs largely through the L1 system. Thus, it is possible that the L2 phonological representations are incomplete, or even incorrect (Darcy et al., 2013), resulting in imprecise representations (akin to the offline argument). This would mean that L2 sound categories (phonology) cannot adequately – or accurately – anchor L2 orthographic representations. L1 correspondences are activated as a sort of fill-in when L2 phonology cannot be activated strongly enough. In this case, the inhibitory effect would operate on phonology rather than orthography, as described above. Further research is required to tease these two explanations apart.

In conclusion, this study adds to the growing literature on the interaction between orthography and phonology in the production of L2 words by showing that spelling influences the production of NAE *schwar* by native Spanish speakers, and that these effects are modulated significantly by input mode. The results suggest online activation of L2 orthography–phonology correspondences, rather than fully integrated representations. Moving forward, research in the field should continue to investigate the link between lexical encoding (perception) and production and orthography.

## Notes

1 Spanish has three letters that are considered separate, individual letters that English does not: <rr>, <ll>, and <ñ>, as in the words <perro> “dog,” <pollo> “chicken,” and <mañana> “tomorrow,” respectively. While <ll> and <rr> occur in words such as “wallet” and “arrive,” they are considered double letters, not one letter alone as in Spanish. In English, both <ll> and <rr> are pronounced as [l] and [r], the same way as their single-letter counterparts <l> and <r>. In Spanish, <rr> and <ll> correspond to different phones, specifically, [r] and in most dialects, [j].

2 There is some disagreement in the literature regarding how best to label these different vowels.

3 There is a large literature on vowel normalization procedures. Since the stimuli were not representative of the full vowel set for each speaker and we were not interested in sociophonetic details, Bark Difference normalization was determined to be adequate. Moreover, Bark Difference also allows clearer comparisons with the traditional F1–F2 vowel plots, facilitating visualization of the data for the Pillai analysis.

## References

- Amengual, M.** (2016). Cross-linguistic influence in the bilingual mental lexicon: Evidence of cognate effects in the phonetic production and processing of a vowel contrast. *Frontiers in Psychology*, 7, 617.
- 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.
- Boersma, P., & Weenink, D.** (2018). Praat: Doing phonetics by computer [Computer program]. Version 6.0.37. Retrieved February, 3, 2018.
- Bradley, Travis G.** (2019). Spanish rhotics and the phonetics-phonology interface. In S. Colina & F. Martínez-Gil (Eds.), *The Routledge Handbook of Spanish Phonology* (pp. 237–258). London/New York: Routledge.
- Bürki, A., Spinelli, E., & Gaskell, M. G.** (2012). A written word is worth a thousand spoken words: The influence of spelling on spoken-word production. *Journal of Memory and Language*, 67(4), 449–467.
- Cutler, A., & Davis, C.** (2012). An orthographic effect in phoneme processing, and its limitations. *Frontiers in Psychology*, 3, 18.
- Darcy, I., Daidone, D., & Kojima, C.** (2013). Asymmetric lexical access and fuzzy lexical representations in second language learners. *The Mental Lexicon*, 8(3), 372–420.
- Darcy, I., Park, H., & Yang, C. L.** (2015). Individual differences in L2 acquisition of English phonology: The relation between cognitive abilities and phonological processing. *Learning and Individual Differences*, 40, 63–72.
- Delattre, P., & Freeman, D.C.** (1968) A dialect study of American r's by x-ray motion picture. *Linguistics*, 6, 29–68.
- Escudero, P., Hayes-Harb, R., & Mitterer, H.** (2008). Novel second-language words and asymmetric lexical access. *Journal of Phonetics*, 36(2), 345–360.
- Escudero, P., & Wanrooij, K.** (2010). The effect of L1 orthography on non-native vowel perception. *Language and Speech*, 53(3), 343–365.
- Fox, R. A., & Jacewicz, E.** (2009). Cross-dialectal variation in formant dynamics of American English vowels. *The Journal of the Acoustical Society of America*, 126, 2603–2618.
- Hagiwara, R.** (1995). Acoustic realizations of American /r/ as produced by women and men. *UCLA Working Papers in Phonetics*, 90, 1–187.
- Harris, J.** (1969). *Spanish phonology*. Cambridge, MA: MIT Press.
- Hay, J., Warren, P., & Drager, K.** (2006). Factors influencing speech perception in the context of a merger-in-progress. *Journal of Phonetics*, 34(4), 458–484.
- Hillenbrand, J., Getty, L. A., Clark, M. J., & Wheeler, K.** (1995). Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America*, 97(5), 3099–3111.
- Hualde, J. I.** (2005). *The sounds of Spanish with audio CD*. New York: Cambridge University Press.
- Kelley, M. C., Tucker, B. V.** (2020). A comparison of four vowel overlap measures. *The Journal of the Acoustical Society of America*, 147(1), 137–145.
- Kuecker, K., Lockenvitz, S., & Müller, N.** (2015). Amount of rhoticity in schwa and in vowel+/r/in American English. *Clinical Linguistics & Phonetics*, 29(8–10), 623–629.
- McAuliffe, M., Socolof, M., Mihuc, S., Wagner, M. & Sonderegger, M.** (2017). Montreal Forced Aligner [Computer program]. Version 1.0.1, retrieved 17 January 2020 from <http://montrealcorpus-tools.github.io/Montreal-Forced-Aligner/>
- McCloy, D.** (2016). *phonR*: tools for phoneticians and phonologists. R package version 1.0-7.
- Mielke, J., Baker, A., & Archangeli, D.** (2016). Individual-level contact limits phonological complexity: Evidence from bunched and retroflex /ɹ/. *Language*, 92(1), 101–140.
- Montant, M., Schön, D., Anton, J.-L., & Ziegler, J. C.** (2011). Orthographic contamination of Broca's area. *Frontiers in Psychology*, 2, 378.
- Morrison, G. S.** (2009). L1-Spanish Speakers' Acquisition of the English/i/—/ɪ/Contrast II: Perception of Vowel Inherent Spectral Change1. *Language and Speech*, 52(4), 437–462.
- Nycz, J., & Hall-Lew, L.** (2013). Best practices in measuring vowel merger. In *Proceedings of Meetings on Acoustics* 166 ASA, 20(1), 060008.
- Pattamadilok, C., Morais, J., Colin, C., & Kolinsky, R.** (2014). Unattentive speech processing is influenced by orthographic knowledge: Evidence from mismatch negativity. *Brain and Language*, 137, 103–111.



- Perre, L., Pattamadilok, C., Montant, M., & Ziegler, J. C. (2009). Orthographic effects in spoken language: On-line activation or phonological restructuring?. *Brain Research*, 1275, 73–80.
- Qu, Q., & Damian, M. F. (2017). Orthographic effects in spoken word recognition: Evidence from Chinese. *Psychonomic Bulletin & Review*, 24(3), 901–906.
- Qu, Q., & Damian, M. F. (2019). Orthographic effects in Mandarin spoken language production. *Memory & Cognition*, 47(2), 326–334.
- Rafat, Y. (2016). Orthography-induced transfer in the production of English-speaking learners of Spanish. *The Language Learning Journal*, 44, 197–213.
- R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>
- van Rij J, Wieling M, Baayen R, van Rijn H (2020). itsadug: Interpreting Time Series and Autocorrelated Data Using GAMMs. R package version 2.4. <https://cran.r-project.org/web/packages/itsadug/index.html>
- Roberto Gonçalves, A., & Silveira, R. (2020). Orthographic effects in speech production: A psycholinguistic study with adult Brazilian-Portuguese English bilinguals. *Revista de Estudos da Linguagem*, 28(3). <http://www.periodicos.letras.ufmg.br/index.php/relin/article/view/16454>
- Roelofs, A. (2006). The influence of spelling on phonological encoding in word reading, object naming, and word generation. *Psychonomic Bulletin and Review*, 13, 33–37.
- Rogers, H. (2014). *The sounds of language: An introduction to phonetics*. New York: Routledge.
- Seidenberg, M. S., & Tanenhaus, M. K. (1979). Orthographic effects on rhyme monitoring. *Journal of Experimental Psychology: Human Learning and Memory*, 5, 546–554.
- Showalter, C. E., & Hayes-Harb, R. (2013). Unfamiliar orthographic information and second language word learning: A novel lexicon study. *Second Language Research*, 29(2), 185–200.
- Solier, C., Perret, C., Baqué, L., & Soum-Favaro, C. (2019). Written training tasks are better than oral training tasks at improving L2 learners' speech production. *Applied Psycholinguistics*, 40(6), 1455–1480.
- Sóskuthy, M. (2017). Generalised additive mixed models for dynamic analysis in linguistics: a practical introduction. *arXiv preprint arXiv:1703.05339*.
- Taft, M. (2006). Orthographically influenced abstract phonological representation: Evidence from non-rhotic speakers. *Journal of psycholinguistic research*, 35(1), 67–78.
- Walker, R., & Proctor, M. (2019). The organisation and structure of rhotics in American English rhymes. *Phonology*, 36(3), 457–495.
- Wieling, M. (2018). Analyzing dynamic phonetic data using generalized additive mixed modeling: a tutorial focusing on articulatory differences between L1 and L2 speakers of English. *Journal of Phonetics*, 70, 86–116.
- Willis, E. W., & Bradley, T. G. (2008). Contrast maintenance of taps and trills in Dominican Spanish: Data and analysis. In *Selected proceedings of the 3rd Conference on Laboratory Approaches to Spanish Phonology* (pp. 87–100). Somerville, MA: Cascadilla Proceedings Project.
- Winter, B., & Weiling, M. (2016). How to analyze linguistic change using mixed models, Growth Curve Analysis and Generalized Additive Modeling. *Journal of Language Evolution*, 1(1), 7–18.
- Wood, S. (2019). mgcv: Mixed GAM Computation Vehicle with Automatic Smoothness Estimation, vsn 1.8-24. <https://cran.r-project.org/web/packages/mgcv/>.
- Young-Scholten, M., & Langer, M. (2015). The role of orthographic input in second language German: Evidence from naturalistic adult learners' production. *Applied Psycholinguistics*, 36(1), 93–114.
- Ziegler, J. C., & Ferrand, L. (1998). Orthography shapes the perception of speech: The consistency effect in auditory word recognition. *Psychonomic Bulletin & Review*, 5, 683–689.

---

Cite this article: Shea, C. (2021). Spanish speakers' English schwa production: Does orthography play a role? *Applied Psycholinguistics*. <https://doi.org/10.1017/S0142716421000266>