Journal of Phonetics 94 (2022) 101180

Contents lists available at ScienceDirect

Journal of Phonetics

journal homepage: www.elsevier.com/locate/Phonetics

Research Article

The impact of orthographic forms on speech production and perception: An artificial vowel-learning study

Antje Stoehr [a,*], Clara D. Martin [a,b]

[a] Basque Center on Cognition, Brain and Language, 20009 Donostia-San Sebastián, Spain
[b] Ikerbasque, Basque Foundation for Science, 48013 Bilbao, Spain

A R T I C L E I N F O

Article history:
Received 22 October 2021
Received in revised form 27 July 2022
Accepted 29 July 2022

Keywords:
Non-native vowel learning
Grapheme-phoneme correspondences
Incongruent orthographic forms
Novel orthographic forms
Speech perception
Speech production

A B S T R A C T

This study investigates the effect of orthographic forms on phonetic aspects of isolated speech sound production and perception. Three groups of 25 L1-Spanish speakers were exposed to /y/ and /ɛ/ in a multi-session learning study. They heard the same vowels presented with: L1-incongruent orthographic forms, novel orthographic forms, or without orthographic forms. After three exposure sessions, participants were tested on vowel production in an elicited production task and vowel perception in a multiple forced choice task. All groups established new /y/ and /ɛ/ production and perception categories. Incongruent orthographic forms led to less target-like category positions for /y/ but not /ɛ/ in production and perception. Novel orthographic forms only facilitated more target-like perception for /y/. In a fourth session, Auditory-only participants were exposed to incongruent orthography for /y/ and novel orthography for /ɛ/. Sequential exposure to incongruent orthography caused less target-like production and perception category positions, while sequential exposure to novel orthography altered neither. Together these results suggest that orthographic forms affect isolated speech sounds and are encoded at the speech sound level. Incongruent grapheme-to-phoneme mappings from L1 to later-learned languages may critically affect the phonetic characteristics of non-native speech sounds, but learning outcomes depend on specific L1-L2 category contrasts.

1. Introduction

It is well established that foreign language (hereafter, L2, irrespective of acquisition order) learners hardly, if ever, achieve a nativelike pronunciation, due to the influence of first language (L1) phonological and phonetic systems (Brennan, Ryan & Dawson, 1975; Ferguson & Garnica, 1975; Flege, 1980, 1981, 1995; Flege, Birdsong, Bialystok, Mack, Sung & Tsukada, 2006; Piske, MacKay, & Flege, 2001; Saito, 2015; Scovel, 1969; Stoehr, Benders, van Hell & Fikkert, 2017). Orthography, which is omnipresent in formal L2 instruction and may also contribute to foreign accentedness in L2, had received surprisingly little attention until recently (for a recent review, see Hayes-Harb & Barrios, 2021). Crucially, mismatches in grapheme-phoneme correspondences (GPCs) are frequent among languages that use the same script. The grapheme <u>, for example, corresponds to /y/ in French but to /u/ in Spanish. If L1-Spanish speakers learn French as an L2, they are confronted with two challenging tasks: learning to perceive and produce /y/ accurately and overcoming the L1 GPC between <u> and /u/. These tasks may be even more challenging when an L2 sound is perceptually close to an L1 sound. L1-Spanish speakers, for instance, find it notoriously difficult to discriminate French or Catalan /ɛ/ from native /e/ in production and perception (Kartushina & Frauenfelder, 2014; Kartushina & Martin, 2019; Pallier, Bosch & Sebastián-Gallés, 1997; Sebastián-Gallés, Echeverría & Bosch, 2005; Sebastián-Gallés, Rodríguez-Fornells, de Diego-Balaguer & Diaz, 2006; Sebastián-Gallés, Vera-Constán, Larsson, Costa & Deco, 2009). The perceptually small difference between /ɛ/ and /e/ may be further reduced by their joint orthographic form <e>.

Despite emerging evidence for orthography-based mispronunciation of L2 words (Bassetti, 2017; Bassetti & Atkinson, 2015; Bassetti, Sokolović-Perović, Mairano & Cerni, 2018; Bürki, Welby, Clément & Spinelli, 2019; Cerni, Bassetti & Masterson, 2019; Nimz & Khattab, 2020; Rafat, 2015; Welby, Spinelli & Bürki, 2021; Young-Scholten & Langer, 2015) it has proved difficult to isolate the effect of orthographic forms

* Corresponding author.
E-mail address: a.stoehr@bcbl.eu (A. Stoehr).

https://doi.org/10.1016/j.wocn.2022.101180
0095-4470/© 2022 The Author(s). Published by Elsevier Ltd.

on pronunciation errors. Classroom learners are frequently exposed to non-native input from peers and teachers, which may reinforce orthography-based mispronunciations. For instance, L1-Spanish learners of French might produce the French word /myl/ <mule> *slipper* more like [mul] due to a GPC mismatch: the grapheme <u> corresponds to /y/ in French and /u/ in Spanish. Alternatively, these effects could arise from frequent exposure to such mispronunciations among non-native French teachers and classmates. In the latter case, auditory exposure to mispronounced forms could have led to the overestimation of orthographic effects in studies on L2 speech sound learning. The present study implements a controlled vowel-learning paradigm that avoids this confound and isolates the effects of orthography on speech sound learning.

1.1. Previous research on orthographic effects in L2 production and perception

Previous studies on the orthographic effects of the L1 on the L2 have focused heavily on word production. These studies have provided strong evidence for orthographically influenced mispronunciations in the L2 when there is an L1-L2 GPC mismatch (Bassetti, 2017; Bassetti & Atkinson, 2015; Bassetti et al., 2018; Bürki et al., 2019; Cerni et al., 2019; Nimz & Khattab, 2020; Rafat, 2015; Welby et al., 2021; Young-Scholten & Langer, 2015). Mismatched L1 and L2 GPCs may lead to less target-like phonetic characteristics, including: vowel formants shifted towards the orthographically-linked L1 vowel (Bürki et al., 2019; Nimz & Khattab, 2020; Rafat, 2015; Welby et al., 2021); production differences between homophonous words with different spellings, such as <sun> versus <son> which share the target pronunciation /sʌn/ (Bassetti & Atkinson, 2015); production of length contrasts between identical vowels or consonants written either as singletons or diagraphs, such as <scene> versus <seen> or <city> versus <kitty> (Bassetti, 2017; Bassetti & Atkinson, 2015; Bassetti et al., 2018; Cerni et al., 2019; Nimz & Khattab, 2020); complete sound substitutions, such as German <Sonne> /zɔnə/ pronounced as [sɔnə] (Rafat, 2015; Young-Scholten & Langer, 2015); and production of silent letters, such as <lamb> /læm/ pronounced as [læmb] (Bassetti & Atkinson, 2015). These effects appear to be robust: they have been observed in L2 speakers with various proficiency levels, ranging from novice learners (Rafat, 2015) to highly proficient L2 speakers (Bassetti et al., 2018) and across learning contexts from lab-based learning (Bürki et al., 2019; Rafat, 2015; Welby et al., 2021) to L2 immersion (Bassetti et al., 2018; Young-Scholten & Langer, 2015). Only when a highly proficient L2 is acquired in early childhood before learning to read, L2 word production appears resistant to mismatched L1-L2 GPCs (Stoehr & Martin, 2022).

Most previous research has investigated orthographic effects on L2 production, with only a small number of studies focused on L2 perception. One of these studies showed that mismatched L1-L2 GPCs negatively affected Spanish-Basque early bilinguals' word and syllable perception (Stoehr & Martin, 2022). This study demonstrated that an orthographic link can be retroactively forged between two distinct L1 and L2 sounds that were both acquired *before* learning to read, showing increased errors in word perception and slower discrimination speeds for meaningless syllables in highly proficient L2 speakers. Other perception studies, primarily dedicated to orthographic effects on novel word learning, have shown mixed results (e.g., Escudero, Simon & Mulak, 2014; Mathieu, 2016; Pytlyk, 2011; Showalter & Hayes-Harb, 2015; Simon et al., 2010). L1-English speakers learning French words containing the vowel contrast /u/ <ou> and /y/ <u> were neither aided nor hindered by the presence of orthographic forms (Simon et al., 2010). A nuanced pattern also emerged when L1-Spanish speakers performed a Dutch vowel discrimination task: the specific GPCs involved, and the difficulty of the sound contrasts accounted for perceptual acuity (Escudero & Wanrooij, 2010). Taken together these results suggest that whether orthography hinders, has no effect on, or facilitates L2 perception, depends on the difficulty of specific L1-L2 sound contrasts and the specific GPCs used. The present study systematically tests if building an association between L1-incongruent orthographic forms and newly learned speech sounds with different levels of perceptual difficulty hinders target-like production and perception.

The impact of orthography on L2 learning may be fundamentally different when L2 and L1 scripts differ, reducing or even eliminating the potential for mismatched L1-L2 GPCs. Several studies have investigated the effect of unknown orthographic forms on L2 word learning and syllable discrimination, showing facilitation (Hayes-Harb & Cheng, 2016; Showalter & Hayes-Harb, 2013), no effect (Hayes-Harb & Hacking, 2015; Pytlyk, 2011; Showalter, 2018; Showalter & Hayes-Harb, 2015), or hindrance (Mathieu, 2016). It appears that learning outcomes are influenced by the difficulty of the sound contrast, the script, and the task. The present study also tests whether learning associations between novel orthographic forms and newly learned speech sounds of varying difficulty facilitates target-like production and perception.

To summarize, previous research has focused almost exclusively on lexical items and suggests that word spelling affects both production and perception. Only two studies have explored the effects of orthography on speech sound perception in the absence of lexical context (Escudero & Wanrooij, 2010; Stoehr and Martin, 2022). These studies provide evidence that distinct orthographic forms may facilitate discrimination whereas mismatched L1-L2 GPCs may harm discrimination. The effect of novel orthographic forms on L2 speech sound learning in either production or perception remains unknown. There is some evidence, moreover, that even when an L2 is acquired before reading acquisition, mismatched L1-L2 GPCs can continue to affect speech perception in adulthood in highly proficient L2 speakers (Stoehr & Martin, 2022). There is no evidence as yet for such retroactive orthographic effects on speech production and perception in learners who are literate in their L1 and who start learning an L2 in the auditory modality before being exposed to orthographic forms.

1.2. Isolating the effect of orthographic forms on L2 speech sound production and perception

The present study implements a controlled speech sound learning paradigm to test whether L1-incongruent (hereafter, conflicting) or novel orthographic forms affect the production and perception of isolated speech sounds in terms of category formation and position. These possible effects are addressed in learning conditions that involve either simultaneous or sequential exposure to orthographic forms. This controlled learning paradigm allows us to isolate the effects of orthographic forms on speech production and perception since the presence and type of orthographic forms is the only task difference between learning groups. Moreover, in contrast to previous work with lexical items, the present study tests the production and perception of single speech sounds to ascertain whether orthographic forms affect pronunciation at the speech sound level as well as at the lexical level, as has been previously reported. If this is the case, it would have crucial implications. First, it would show that the mismatch between L1-L2 GPCs may predict L2 learners' pronunciation. This would have important implications for L2 speech sound production and perception models, which do not yet incorporate orthography as a factor influencing L2 acquisition outcomes. Second, if speech sounds are affected by orthographic forms, it may suggest that previously found orthographic effects at the lexical level are not exclusively driven by co-activation of L1 orthographic forms during L2 word processing. Instead, orthographically influenced mispronunciations of L2 words may also be driven by distorted L2 speech sound representations.

1.3. Models on L2 speech production and perception

The Speech Learning Model (SLM; Flege, 1995) and the Perceptual Assimilation Model for L2 learning (PAM-L2; Best & Tyler, 2007) provide crucial theoretical background for the present study. Although the SLM and PAM-L2 differ in many respects, both models assume that L2 speakers have a common phonetic space that accommodates both L1 and L2 categories. This space remains somewhat flexible throughout life, enabling the formation of new categories. If a new sound is perceived to be sufficiently different from the learner's existing vowel categories, the SLM and PAM-L2 both predict a new category will be formed, although it may differ from the category a native speaker of the target language has for that same sound. The formation of new categories depends on how distinct newly encountered sounds are perceived to be from existing categories: If a new sound is not perceived to be sufficiently different from the closest existing L1 category, it is likely to be classified as belonging to this already existing sound category and no new category will be formed. This process is known as *equivalence classification* in the SLM framework and as *perceptual assimilation* in the PAM-L2 framework.

1.4. The present study

In the present study, L1-Spanish speakers learned the close anterior rounded vowel /y/ and the open-mid anterior unrounded vowel /ɛ/ during three exposure sessions. Both sounds are reportedly difficult for Spanish speakers but to a different extent. /y/ falls within an unoccupied area of the Spanish vowel space, and Spanish leaners of French can produce /y/ with formants that differ from those for L1 vowels, although their pronunciation is still not nativelike (Racine & Detey, 2019). /ɛ/ may be considered an allophone of /e/, although evidence is debatable (see, e.g., de Anda, 2013; Morrison, 2004; Navarro Tomás, 1918, 1965). Importantly, /ɛ/ falls within the vowel space occupied by Spanish /e/ (Meunier, Frenck-Mestre, Lelekov-Boissard & Le Besnerais, 2003), and L1-Spanish learners of French or Catalan find it notoriously difficult to distinguish /ɛ/ from /e/ in production and perception (Kartushina & Frauenfelder, 2014; Kartushina & Martin, 2019; Pallier et al., 1997; Sebastián-Gallés et al., 2005, 2006, 2009).

According to the predictions made by both the SLM and PAM-L2, the non-native vowel /y/ might form its own category if it is perceived to be sufficiently different from the closest Spanish native categories /i/ (/y/ is rounded while /i/ is unrounded) and /u/ (/y/ is anterior [=front] while /u/ is posterior [=back]). If so, /y/ would count as new phone in SLM terms and as uncategorized sound using PAM-L2 terminology. In contrast, the non-native vowel /ɛ/ might be subject to equivalence classification/perceptual assimilation, since /ɛ/ is reportedly difficult for L1-Spanish speakers to discriminate from native /e/ in both production and perception (Kartushina & Frauenfelder, 2014; Kartushina & Martin, 2019; Pallier et al., 1997; Sebastián-Gallés et al., 2005, 2006, 2009). Neither the SLM nor the PAM-L2 make specific predictions about the role of orthography, although the PAM-L2 acknowledges that a shared orthographic link between two similar L1 and L2 sounds may contribute to perceptual assimilation. Given the predictions of these two models and the research into orthographic effects on production and perception reported above, the present study investigates whether conflicting and novel orthographic forms influence the learning of vowel categories and category positions in production and perception.

To address these questions, participants were divided into two experimental groups, plus a control group to compare performance. The Conflicting Orthography group simultaneously heard the vowels and saw the respective L1 graphemes, <u> and <e>, on screen; the Novel Orthography group simultaneously heard the vowels and saw novel orthographic symbols on screen. The control group, Auditory-only, heard the vowels but did not receive any visual input. After the third exposure session, participants were tested on their production and perception of the newly learned vowels. Our precise predictions regarding category formation of /y/ and /ɛ/ sounds are delineated below.

Regarding /y/, we predicted that participants in all groups would form new /y/ categories in production and perception given the considerable acoustic differences between /y/ and the surrounding native vowels /i/ and /u/. However, we predicted the position of this category would differ depending on whether GPCs were conflicting or novel. Specifically, we predicted that the Conflicting Orthography group would form production and perception categories for /y/ at a more posterior position, corresponding to smaller F2 values, compared to the Auditory-only control group. This would be due to interference from the posterior L1 vowel /u/, which shares the grapheme <u> with the newly learned vowel /y/. We further hypothesized that the Novel Orthography group would benefit from the link to a novel grapheme, which may reinforce the distinction between /y/ and L1 vowels. If so, this would result in more precise production and perception of /y/ with more anterior production and perception categories, corresponding to greater F2 values, compared to the Auditory-only control group.

For /ɛ/, we predicted that the Conflicting Orthography group would be strongly influenced by the orthographic link <e> between /ɛ/ and Spanish /e/, which we hypothesized would lead participants in this group to map /ɛ/ onto /e/, blocking new category formation. Consequently, we predicted production and perception categories to be at a higher position, with smaller F1 values, than those of the Auditory-only control group. We further predicted that the Novel Orthography group would be able to acquire a distinct category for /ɛ/ in production and perception because the link to a novel grapheme would help participants distinguish /ɛ/ from the acoustically similar native vowel /e/. We predicted that the link to a novel grapheme would, moreover, help the Novel Orthography group to produce and perceive /ɛ/ more precisely, leading to lower category positions in production and perception, with greater F1 values, compared to the Auditory-only control group. These hypotheses are summarized in Table 1.

Participants in the Auditory-only (control) group returned for a fourth session, to assess whether sequential exposure to conflicting and novel orthographic forms modulates learning outcomes in production and perception. They were exposed to a conflicting GPC for /y/ and novel orthography for /ɛ/. The exposure to conflicting and novel orthography was not counterbalanced because, as explained above, we predicted the largest impact of conflicting orthography on /y/ and the largest impact of novel orthography on /ɛ/. Subsequently, these participants were retested on vowel production and perception.

We hypothesized that sequential exposure to conflicting orthographic forms during the fourth session would bring production and perception categories for /y/ to a more posterior position, with smaller F2 values, relative to the production and perception tests performed earlier without orthographic input. This could be attributed to the influence of the L1 vowel /u/, with which /y/ shares the grapheme <u>, in the fourth session. In addition, we predicted that sequential exposure to novel orthographic forms would facilitate target-like learning. If this is the case, production and perception categories for /ɛ/ should move to a lower position, corresponding to greater F1 values, compared to the earlier auditory-only production and perception tests.

2. Method

2.1. Participants

Seventy-five female L1 speakers of Spanish (M_{age} = 22 years, SD_{age} = 2.6 years, $range_{age}$ = 18–31 years) were divided into three groups of 25. All participants had grown up in Spanish-speaking households and reported some knowledge of Basque and English, which are compulsory languages in the school system in the Basque Autonomous Community in Spain where the study was conducted. Knowledge of Basque and English were not expected to influence the results of the present study for two reasons: First, Spanish and Basque are identical in terms of vowel inventory and vowel spelling; second, participants' relatively low English skills (Table 2) and lack of exposure to native English speaker input made it unlikely that they had full command of English phonology. Nevertheless, to account for any possible effects of Basque or English, the groups were matched on Basque and English proficiency (Table 2). Participants were recruited from the subject pool of the Basque Center on Cognition, Brain and Language (BCBL), which allowed us to match the three groups prior to testing on age, education, IQ (Kaufman Brief Intelligence Test; Kaufman & Kaufman, 2004), and both Basque and English age of acquisition and proficiency as measured by the Basque, English and Spanish test (BEST; de Bruin, Carreiras & Duñabeitia, 2017). The BEST tests Basque, English, and Spanish proficiency using picture naming tasks comprising 65 non-cognate words per language, semi-structured interviews conducted by linguists, and the LexTALE test of vocabulary size (Lemhöfer & Broersma, 2012). As part of the present study, participants were additionally tested on their reading skills (Lallier, Galparsoro, Redondo & Carreiras, 2022) and inhibitory skills (Darcy, Mora & Daidone, 2016; Lev-Ari & Peperkamp, 2013) to ensure that the groups did not differ on these potentially important variables (see Appendix A for more details on these measures). All variables matched across groups are presented in Table 2. Another nine participants were tested but were excluded from the analyses due to technical problems resulting in data loss (n = 3), experimenter error (n = 2), inability to complete all experimental sessions (n = 2), or exposure to a fourth language (n = 2). Participants received monetary compensation for their time and gave written, informed consent prior to starting the experiments. The experiments had previously been approved by the BCBL Ethics Committee.

2.2. Stimuli

2.2.1. Exposure stimuli

During the exposure phase, participants listened to 18 tokens each of /y/ and /ɛ/. Six female L1 speakers of French, hereafter *talkers*, recorded the isolated vowels. Talkers 1–3 each provided six /y/ stimuli and talkers 4–6 each provided 6 /ɛ/ stimuli. Talkers for each vowel were selected based on the distribution of their production to ensure that the tokens for each vowel formed a unimodal distribution (Fig. 1). Table 3 displays the mean formant values and vowel durations by talker and vowel.

Table 1
Hypotheses for production and perception outcomes by group and vowel relative to the Auditory-only control group.

		Conflicting orthography	Novel orthography
/y/	Acquisition of new category	yes	yes
	Category position	influenced by /u/ → posterior (smaller F2)	precise → anterior (greater F2)
/ɛ/	Acquisition of new category	no → merged with /e/	yes
	Category position	influenced by /e/ → higher (=smaller F1)	precise → lower (=greater F1)

Table 2
Participant-specific variables by group.

		Conflicting Orthography	Novel Orthography	Auditory-only	*p*
	Age (years)	22.32 (2.34)	21.88 (3.06)	22.08 (3.11)	0.862
	Education[1]	2.56 (0.77)	2.48 (1.05)	2.17 (0.96)	0.304
	Verbal IQ	101.6 (9.39)	99.28 (9.73)	102.17 (7.43)	0.488
	Matrix IQ	109.92 (8.16)	108.48 (9.73)	108.83 (9.69)	0.848
	Combined IQ	104.4 (8.86)	102.12 (9.16)	104.29 (8.30)	0.590
	Reading skills	225/278 (31.23)	221/278 (30.39)	219/278 (32.86)	0.772
	Inhibitory skills	1.03 (0.52)	0.96 (0.47)	0.84 (0.47)	0.352
Basque	Interview[2]	3.24/5 (1.20)	3.32/5 (1.55)	3.16/5 (0.99)	0.903
	Picture naming	37.52/65 (19.44)	39.36/65 (22.40)	39.68/65 (15.78)	0.914
	LexTALE	79.36% (13.51)	77.92% (17.04)	79.48% (12.70)	0.915
	AoA (years)	2.76 (1.67)	2.68 (4.47)	3.28 (3.36)	0.828
English	Interview	3.2/5 (0.82)	2.84/5 (0.78)	2.92/5 (0.70)	0.201
	Picture naming	40.52/65 (10.26)	39.8/65 (11.85)	41.16/65 (10.38)	0.907
	LexTALE	62.9% (11.29)	63.75% (5.51)	66.2% (11.37)	0.485
	AoA (years)	5.04 (1.90)	5.56 (2.50)	5.68 (2.27)	0.564

Education: 1 = high school; 2 = professional training; 3 = university; 4 = postgraduate degree.
[2] Interview scores: 1 = speakers are hardly able to communicate in the tested language and mostly produce isolated words or memorized forms. 2 = speakers can communicate in simple sentences with frequent pauses and errors in the tested language, which may make them difficult to understand. 3 = speakers are fluent in the tested language, able to speak at length with a wide range of vocabulary and are generally easy to understand though they make some mistakes. 4 = speakers are highly fluent in the tested language and are able to talk about a wide range of topics, with occasional errors during long and difficult sentences. 5 = native speaker competence.

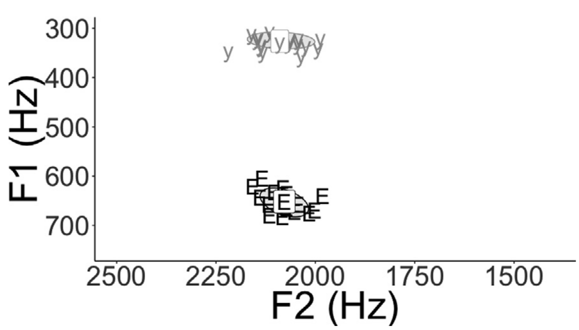

Fig. 1. Exposure vowel tokens in the F1-F2 space.

2.2.2. Perception stimuli

Stimuli for the vowel perception tasks were synthesized in Praat software (Boersma & Weenink, 2017) using a fundamental frequency of 120 Hz. Vowel tokens were sampled from the F1-F2 plane in psychoacoustically equidistant steps along the Mel scale, following a similar procedure as described in Chládková and Escudero (2012). The F1 was sampled in 32 steps from 200 Hz to 1000 Hz, and the F2 was sampled in 24 steps ranging from 500 Hz to 3000 Hz. Combinations of F1 and F2 with F1 equal to or higher than F2 were excluded. Tokens with very high F1 and F2 values that fell outside the range of the possible human vowel space were also excluded. This procedure resulted in 613 unique F1-F2 vowel tokens (Fig. 2). Tokens with F2 below 1500 Hz had F3 set to 2500 Hz; tokens with F2 higher than 1500 Hz had F3 set to 1000 Hz higher than F2. All formants were steady-state throughout vowel production. Vowel duration was 200 ms.

All exposure and perception stimuli were scaled to 70 dB, digitized at 44,100 Hz, and saved as.wav files. Fifty milliseconds of silence were added to the beginning of each audio file to allow for sufficient loading time in the experimental software without any risk of losing auditory information.

2.3. Apparatus and procedure

Participants were tested individually in sound-attenuating chambers at the BCBL satellite laboratory at the University of the Basque Country in Donostia-San Sebastián. The experiments were run on a desktop computer. Open Sesame software (version 3.2.8; Mathôt, Schreij, & Theeuwes, 2012) was used for the exposure phases, the speech production test, and the inhibition and reading control tasks; Praat software (Boersma & Weenink, 2017) was used for the perception test. Stimuli were presented binaurally over Sennheiser GSP 350 headphones.

Participants came to the laboratory on three separate days within an average period of 4 days (range 3–8 days) to allow for learning benefits associated with sleep consolidation (Davis, Di Betta, Macdonald & Gaskell, 2009; Dumay & Gaskell, 2007). During the first session, participants were initially tested on L1 vowel production and perception. They then completed the first exposure phase to the non-native vowels. During the second session, participants completed the second

Table 3
Mean formant values and duration of the exposure vowel tokens by talker.

Vowel	Talker	Number of tokens	F1 in Hz (SD)	F2 in Hz (SD)	Duration in ms (SD)
[y]	1	6	328 (15)	2142 (50)	236 (14)
	2	6	327 (8)	2026 (31)	270 (18)
	3	6	323 (17)	2101 (43)	213 (11)
[ɛ]	4	6	650 (26)	2065 (54)	219 (18)
	5	6	650 (22)	2094 (25)	171 (18)
	6	6	655 (22)	2076 (60)	233 (12)

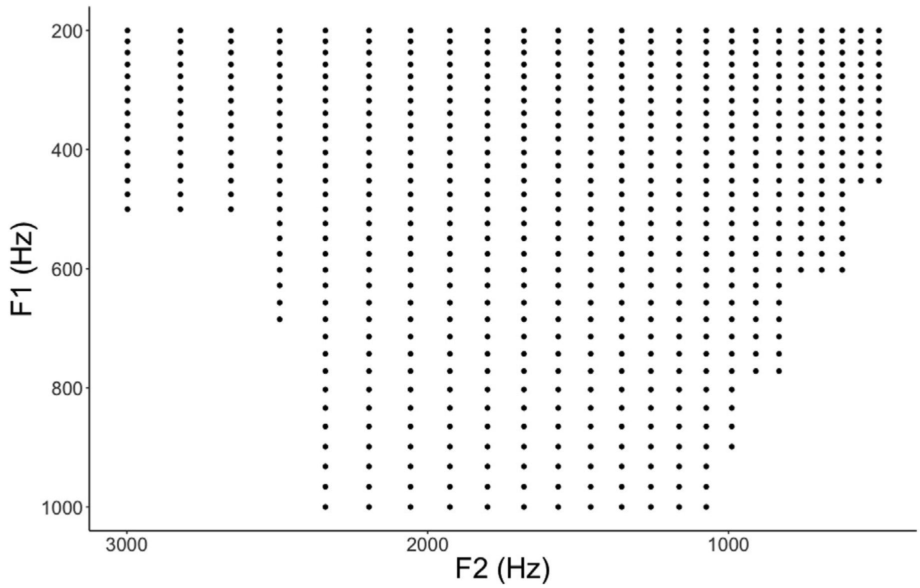

Fig. 2. Distribution of the perception stimuli in the F1-F2 space.

exposure phase. Afterwards, they were tested on their inhibitory skills using the Spanish adaptation of Lev-Ari and Peperkamp's (2013) retrieval-induced inhibition task (Darcy et al., 2016). They then performed the Alondra reading test (Lallier et al., 2022), which is the Spanish adaptation of the original French Alouette reading test (Lefavrais, 1967, 2005). Both tasks are described in detail in Appendix A. The third session started with the third exposure phase, followed by production and perception tests of the newly learned vowels. Afterwards, participants were re-tested on their production and perception of their native vowels for a different project. The first and second sessions each lasted 1.5 h and the third session lasted 2 h. Participants in the Auditory-only group returned to the laboratory for a fourth session on average two days after the third session (range 1–6 days). During the fourth session, they were exposed to orthographic forms during an additional exposure phase. This was followed by a test of their production and perception of the newly learned vowels. Afterwards, they completed L1 production and perception tests for a different project. The fourth session lasted approximately 2.5 h.

2.3.1. Exposure

During all sessions, the exposure phases were blocked. Half of the participants in each group started with /y/, and the other half started with /ɛ/. During exposure, participants listened to a sequence of three different tokens of the same vowel with an interstimulus interval of 1000 ms. These three vowel tokens were randomly selected from the pool of 18 tokens for each vowel. During the auditory presentation of the vowels, their orthographic forms were shown on screen. For the Conflicting Orthography group, /y/ was represented as <u> and /ɛ/ as <e>. In Spanish the grapheme <u> corresponds to /u/ and <e> to /e/, thus, the Conflicting Orthography group was faced with mismatched GPCs for the newly learned vowels and L1. For the Novel Orthography group, the orthographic form of /y/ was <ɸ> and /ɛ/ was represented by <ɘ>. The Auditory-only control group did not view any orthographic forms. The last vowel token in the triad was followed by 2000 ms of silence, after which an auditory beep – and additionally in the case of the Conflicting and Novel Orthography groups, the orthographic symbol turning red – prompted the participants to produce the vowel. At the moment these cues were presented, the recording was initiated, and remained active for 1500 ms. After an intertrial interval of 1000 ms, the next trial started. We opted to include a production component in the exposure phase for two reasons. First, to mimic more natural learning conditions in which language learners generally start to produce sounds and words from the initial learning stages; second, to make the exposure phase more engaging. During the first and third sessions, the exposure phase consisted of 150 trials each including three vowel variants, giving exposure to 450 tokens per vowel, and lasted 18.5 min. During the second session and the fourth session (the latter administered exclusively to the Auditory-only group), the exposure phase for each vowel consisted of 300 trials each including three vowel variants, for a total exposure to 900 tokens per vowel, and lasted 37 min. To ensure that participants were paying attention to the orthographic forms provided on the screen during exposure, 10% of the trials contained catch trials, during which an asterisk shortly appeared on the screen. Participants were instructed to press the space bar whenever the asterisk appeared. If participants failed to do so, a buzz-sound was played at the end of the trial. These catch trials were also administered to the auditory-only group.

2.3.2. Vowel production

Each participant completed separate vowel production tasks both in their L1 and with the newly learned vowels. The L1 production tasks were administered during the first session before exposure to the non-native vowels, and during the third and fourth (Auditory-only group) sessions after the non-native vowel production task. All five native Spanish vowels /i/, /e/, /a/, /o/, /u/ were included in the L1 tasks. In this study, the results for /i/, /e/ and /u/ elicited in the first session are reported; /a/ and /o/ were elicited for a different study. The vowels were elicited in vowel-specific blocks, which were presented in randomized order. In contrast to the exposure phases, no auditory vowel exemplars were provided during the production test. Instead, the orthographic form of the vowel appeared on screen, and participants were instructed to produce the vowel as soon as the orthographic form turned red, which occurred after a jittered interval ranging from 1500 to 2000 ms. At the moment that the orthographic form turned red, an auditory beep was also played, and the recording was initiated. The recorder remained active for 2000 ms. Afterwards, a blank screen appeared for 500 ms before the next trial started. Each of the five vowel blocks contained 25 trials, resulting in a total of 125 productions per participant.

The non-native vowel production task was administered in the third and fourth (Auditory-only group) sessions before the second and third L1 production task. At the beginning of the non-native vowel production task, participants listened to three tokens of the vowel they would be prompted to produce. The Conflicting and Novel Orthography groups then followed the same procedures as in the L1 production task with the orthographic form of the newly learned vowel on screen. The Auditory-only group received no orthographic input as only a fixation cross appeared on screen. In each group, half of the participants produced /y/ first while the other half of the participants first produced /ɛ/. Each vowel was produced 25 times, resulting in a total of 50 productions by each participant.

2.3.3. Vowel perception

Vowel perception was tested using a multiple forced choice task. The L1 vowel perception task was administered during the first session and again in the third and fourth (Auditory-only group) sessions, after the non-native vowel perception task. For L1 vowel perception, participants saw the orthographic forms of the five Spanish vowels <a>, <e>, <i>, <o>, <u> on screen. In addition, participants were given the option of responding “none of the choices” (*ninguna de las opciones*). Participants listened to one vowel at a time and then had to click on the corresponding response option on the screen. The next trial began 500 ms after participants gave their response. After every 62 trials, participants were given the option to take a short break, resulting in 9 possible breaks before they had categorized all 613 vowel tokens. The L1 perception task lasted approximately 30 min. The experiment was preceded by a practice phase, which had been set up in the same way as the main experiment. The practice phase consisted of 20 trials that were randomly selected from the 613 experimental stimuli for each participant.

The non-native vowel perception task was created in line with the L1 vowel perception task. Participants listened to the same 613 vowel tokens, but they were only given three response options. All groups had the option to respond “none of the choices”. In addition, the Conflicting Orthography group had <u> and <e> as response options, while the Novel Orthography group saw the response options <ɸ> and <ɞ>. The Auditory-only group was given the response options “sound 1” (*sonido 1*) and “sound 2” (*sonido 2*), corresponding to the order in which they had learned the sounds across all three exposure sessions. Prior to starting the non-native vowel perception task, the experimenter ensured that the participants in the Auditory-only group remembered that learning order. As in the L1 vowel perception task, 20 practice trials were included.

2.4. Acoustic measurements

All 10 625 vowels produced in this study were labeled using the TextGrid (silences) function in Praat software (Boersma & Weenink, 2017). All automatically inserted intervals were checked by the first author. In this step, noisy productions were labeled so they could later be excluded from the analyses. For each vowel token, F1, F2, and F3 were measured using a modified version of the Optimized Formant Ceiling algorithm proposed by Escudero, Boersma, Schurt Rauber and Bion (2009) to account for individual differences in vocal tract length. Instead of using the default value of 5500 Hz as a maximum ceiling for female speakers, formants were measured at various maximum ceilings, ranging from 4500 Hz to 6500 Hz in steps of 100 Hz. The optimal ceiling for a speaker was determined as follows: First, the standard deviation for each formant of each vowel was calculated at each measurement point. Then, the standard deviations for F1, F2, and F3 for each vowel were combined at each measurement point. Next, these combined standard deviations were averaged across all L1 and newly learned vowels measured at the same ceiling for each speaker. The ceiling that yielded the lowest mean standard deviation for all vowels combined was selected as the optimal ceiling for a speaker. The mean ceilings per group were 5800 Hz (*SD* 405 Hz) for the Conflicting Orthography group, 5735 Hz (*SD* 512 Hz) for the Novel Orthography group, and 5684 Hz (*SD* 456 Hz) for the Auditory-only group.

3. Results

This section presents two sets of main analyses: first, analyses to establish whether orthographic forms during exposure influenced the formation (relative to L1) and position of new vowel categories in both production and perception; second, analyses to test for any effects of sequential exposure to first auditory-only then auditory-orthographic input.

Data were analyzed for vowel formants with linear mixed-effects models using the lme4 package version 1.1–29 (Bates, Mächler, Bolker & Walker, 2015) in RStudio version IDE (RStudio Team, 2022) run on R version 4.2.0 (R Core Team, 2022). The *p*-values for *t*-statistics were obtained using the lmerTest package version 3.1–3 (Kuznetsova, Brockhoff, & Christensen, 2017). Data were visualized using the ggplot2 package version 3.3.6 (Wickham, 2016). Significant findings are presented below; complete model outputs are available in the Appendix (Tables B1-B8).

In total, 99.2% of the production data (10 540 productions) were included in analyses. Eighty-five productions were excluded because of noisy recordings, for example, due to an echo in the microphone or participant-related noise, such as coughing or yawning. The perception data comprised 19 952 tokens that were categorized as either /y/ (6983) or /ɛ/ (12 969). Table 4 shows the distribution of categorizations by group and vowel. One participant in the Novel Orthography group and one participant in the Auditory-only group were excluded from perception analyses because they confused the response options during the categorization task.

Table 4
Number of tokens categorized as /y/ and /ɛ/ by group.

Group	Vowel	Number of categorizations
Conflicting Orthography	/y/	2495
	/ɛ/	4575
Novel Orthography	/y/	2073
	/ɛ/	4208
Auditory-only	/y/	2415
	/ɛ/	4186

3.1. Orthographic effects on non-native vowel learning and category position

3.1.1. The vowel /y/

Given the acoustic differences between /y/ and the surrounding native vowels /i/ and /u/, we predicted that participants in all groups would produce and perceive /y/ to be distinct from L1 vowel categories. We further hypothesized that the Conflicting Orthography group would produce a less target-like /y/ with smaller (more posterior) F2 values and perceive tokens with smaller F2 values to be /y/ in contrast to the Auditory-only control group. This could be attributed to interference from L1 /u/, which shares the grapheme <u> with the newly learned vowel. In addition, we predicted that the link to a novel grapheme would help the Novel Orthography group produce and perceive /y/ more accurately, leading to greater (more anterior) F2 compared to the Auditory-only control group.

To test whether the groups produced and perceived /y/ distinctly from surrounding native vowels, one linear mixed-effects model was conducted per modality. The production model evaluated the F2 of /y/ relative to the F2 of /i/ and /u/. The perception model evaluated the F2 of tokens categorized as /y/ relative to the F2 of tokens categorized as /i/ or /u/. For the three-level variables Group and Vowel, we used deviation coding to create two contrasts of interest each. For Group, the first contrast compared the Conflicting Orthography group [coded as 1] to the Auditory-only control group [coded as −1], and the second contrast compared the Novel Orthography group [coded as 1] to the Auditory-only control group [coded as −1]. For Vowel, the first contrast compared /y/ [coded as 1] to /i/ [coded as −1], and the second contrast compared /y/ [coded as 1] to /u/ [coded as −1]. The production and perception models had *Vowel* and *Group* as fixed effects including an interaction term. Random intercepts for *Participant* and by-Participant random slopes for *Vowel* were included (model R code: F2 ~ Vowel * Group + (1 + Vowel | Participant)).

The production model showed that /y/ was produced with a smaller F2 than /i/, representing a more posterior position ($\beta = -791.430$ $SE = 20.150$, $t = -39.280$, $p < 0.001$); /y/ productions also exhibited a greater F2 than /u/, representing a more anterior position ($\beta = 1096.030$, $SE = 27.750$, $t = 39.490$, $p < 0.001$). The model, furthermore, detected a significant difference between the Conflicting Orthography group and the Auditory-only group ($\beta = -70.710$, $SE = 26.770$, $t = -2.641$, $p = 0.010$) and a significant interaction between Group (Conflicting Orthography group vs Auditory-only group) and Vowel (/y/ vs /i/; $\beta = -74.130$, $SE = 28.490$, $t = -2.601$, $p = 0.011$). The main effect of Group and the interaction between Group and Vowel show that the Conflicting Orthography group and the Auditory-only group differed in terms of the F2 for /y/ which the Conflicting Orthography group produced with a smaller F2, at a more posterior position, than the Auditory-only group, indicating orthographic interference effects. The Novel Orthography group did not detectably differ from the Auditory-only group in /y/ production; no other significant effects or interactions were observed (Appendix, Table B1).

Similarly, the perception model detected that tokens categorized as /y/ had a smaller F2, corresponding to a more posterior position, than tokens categorized as /i/ ($\beta = -695.100$, $SE = 20.200$, $t = -34.413$, $p < 0.001$), and a greater F2, corresponding to a more anterior position, than tokens categorized as /u/ ($\beta = 829.600$ $SE = 20.000$, $t = 41.487$, $p < 0.001$). A significant difference between the Conflicting Orthography group and the Auditory-only group ($\beta = -81.830$, $SE = 23.380$, $t = -3.500$, $p < 0.001$), and significant interactions between Group (Conflicting Orthography group vs Auditory-only group) and Vowel (/y/ vs /i/; $\beta = -112.410$, $SE = 28.480$, $t = -3.948$, $p < 0.001$; /y/ vs /u/; $\beta = -64.070$, $SE = 28.130$, $t = -2.278$, $p = 0.026$) indicate that the Conflicting Orthography group categorized tokens with a smaller F2, at a more posterior position, as /y/ than the Auditory-only control group, but the two groups did not differ in terms of F2 on tokens categorized as /i/ or /u/. Moreover, a significant difference between the Novel Orthography Group and the Auditory-only group ($\beta = 61.610$, $SE = 23.650$, $t = 2.605$, $p = 0.011$), and a significant interaction between Group (Novel Orthography group vs Auditory-only group) and Vowel (/y/ vs /u/; $\beta = 74.690$, $SE = 28.440$, $t = 2.626$, $p = 0.011$) demonstrate that the Novel Orthography group categorized tokens with greater F2, at a more anterior position, as /y/ than the Auditory-only control group, but the two groups did not detectably differ in terms of F2 on tokens categorized as /u/. No significant interaction was observed between Group (Novel Orthography group vs Auditory-only group) and Vowel (/y/ vs /i/; $\beta = 55.280$, $SE = 28.700$, $t = 1.926$, $p = 0.058$; Appendix, Table B2).

Taken together, these results demonstrate that all groups acquired new /y/ categories with distinct F2 from their L1 /i/- and /u/-categories in production and perception (Fig. 3). Moreover, the presence of conflicting orthographic forms during exposure caused more posterior (less target-like) category positions in production and perception, whereas the presence of novel orthographic forms led to a more anterior – and thus target-like – position for /y/ in perception but not in production.

3.1.2. The vowel /ɛ/

We predicted that the presence of Conflicting Orthography would hinder distinct production and perception of /ɛ/ and /e/, leading to production and perception of smaller F1 values corresponding to a higher position for /ɛ/ compared to the Auditory-only control group. Conversely, we predicted the association between a novel grapheme and /ɛ/ would facilitate learning, causing participants in the Novel Orthography group to produce and perceive /ɛ/ most distinctly from /e/. Consequently, we predicted the Novel Orthography group would produce and perceive /ɛ/ with greater F1, corresponding to a lower position, than the Auditory-only control group.

To test whether the groups acquired a new category for /ɛ/ in production and perception, one linear mixed-effects model per modality was carried out. The production model had F1 pro-

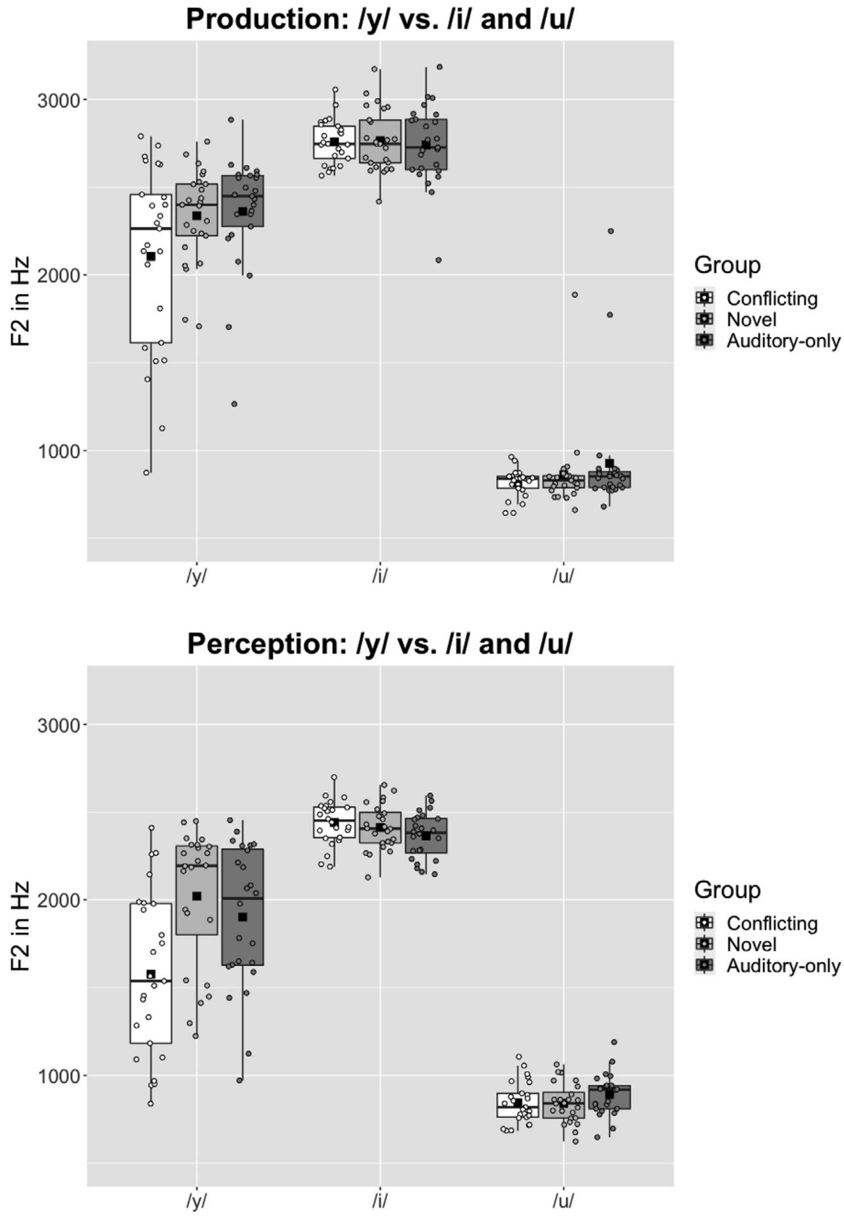

Fig. 3. /y/ production (top) and /y/ perception (bottom). The boxes show the first and third quartiles, the line indicates the median, the square shows the mean, and each dot represents an individual participant.

duction as dependent variable while the perception model had the F1 of tokens categorized as /ɛ/ and /e/ as dependent variable. As fixed effects, both models had *Vowel* (deviation coded as /e/ −1; /ɛ/ 1) and *Group* (using deviation coding to compare the Conflicting Orthography group [coded as 1] to the Auditory-only control group [coded as −1], and the Novel Orthography group [coded as 1] to the Auditory-only control group [coded as −1]). An interaction term between Vowel and Group was included. The model had random intercepts for *Participant* and by-Participant random slopes for *Vowel* (model R code: F1 ~ Vowel * Group + (1 + Vowel | Participant)).

The production model showed that /ɛ/ was produced with greater F1, and, therefore, at a lower position, than /e/ (β = 45.213, *SE* = 5.091, *t* = 8.881, $p < 0.001$). No other significant effects or interactions were observed (Appendix; Table B3). In line with production, the perception model detected that tokens categorized as /ɛ/ had greater F1 than tokens categorized as /e/ (β = 16.140, *SE* = 2.978, *t* = 5.420, $p < 0.001$). No other significant effects or interactions were observed (Appendix, Table B4). The combined production and perception results show that against our prediction and despite the similarity of non-native /ɛ/ and L1 /e/, participants were able to produce and perceive these vowels distinctly, independent of their exposure conditions. Fig. 4 shows the F1 of /ɛ/ and /e/ in production and perception by group.

3.2. Sequential exposure to orthographic forms

Participants in the Auditory-only group were tested on their vowel production and perception twice. The first test occurred during the third session, after three sessions in which they had been exposed to the vowels without any visual input, serving as controls for the orthographic groups. During the fourth session, they were exposed to conflicting orthography for /y/ (<u>)

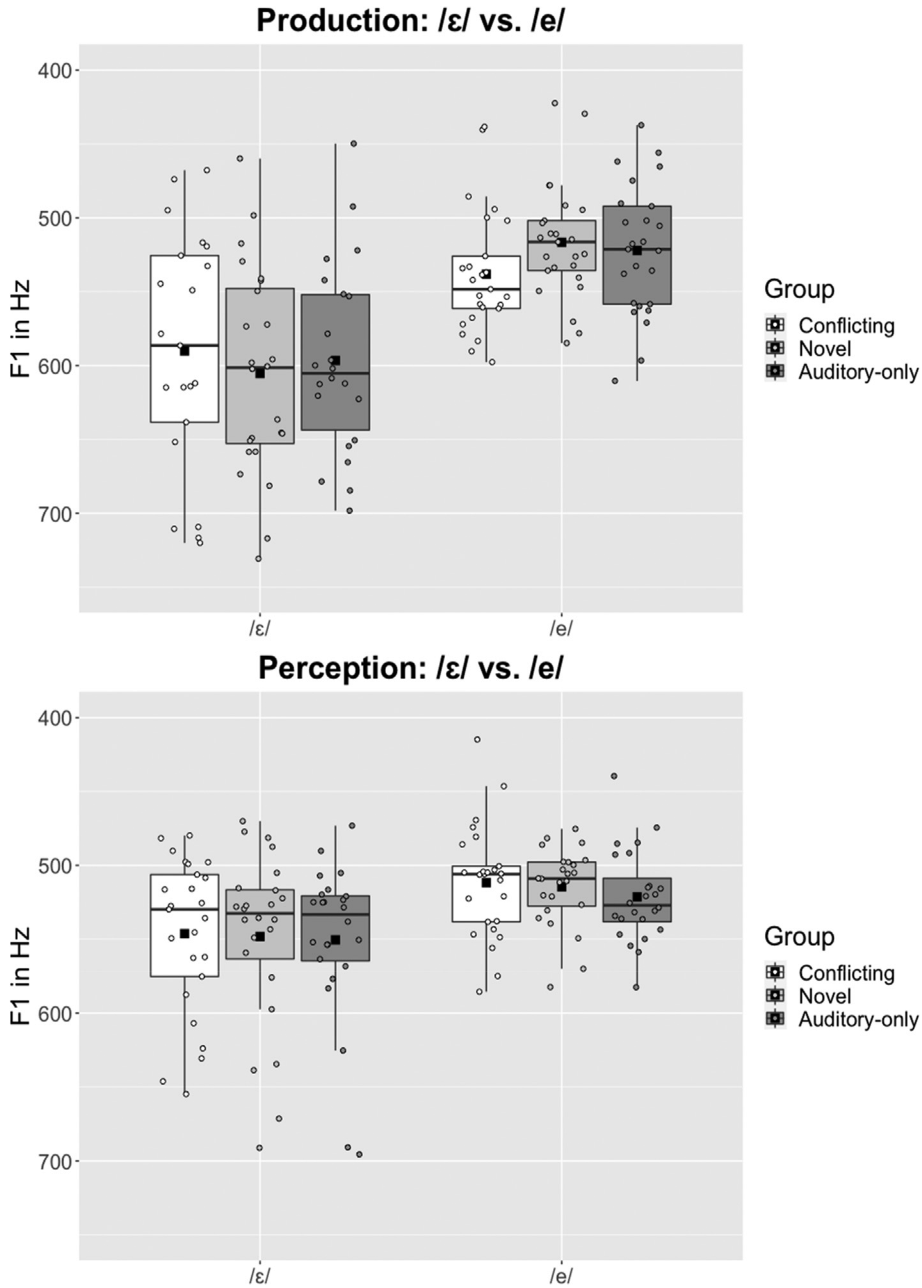

Fig. 4. /ɛ/ production (top) and /ɛ/ perception (bottom). The boxes show the first and third quartiles, the line indicates the median, the square shows the mean, and each dot represents an individual participant.

and novel orthography for /ɛ/ (<ɞ>). If exposure to conflicting orthography hinders learning, we hypothesized that /y/ production and perception categories would move to a more posterior position corresponding to smaller F2 values relative to the first test due to the influence of the L1 vowel /u/, with which /y/ now shared the grapheme <u>. If exposure to novel orthography facilitates target-like learning, we hypothesized that /ɛ/ production and perception categories should move to a lower position corresponding to greater F1 values relative to the first test due to dissociation from the L1 vowel /e/ caused by the novel grapheme <ɞ>. In the following, we present models of F2 changes in /y/ production and perception and F1 changes in /ɛ/ production and perception after sequential exposure to orthographic forms.

3.2.1. Production and perception of /y/

Two linear mixed-effects models were conducted. The production model had F2 production as the dependent variable and the perception model had the F2 of tokens categorized as /y/ as the dependent variable. Both models included *Time* (session 3 [auditory-only] deviation coded as -1, session 4 [orthographic input] deviation coded as 1) as fixed effect and random intercepts for *Participant* with by-Participant random slopes for *Time* (model R code: F2 $\sim$ Time + (1 + Time | Participant)).

The production model detected smaller F2 values at the second test, showing that /y/ was produced at a more posterior position after introducing conflicting orthography ($\beta = -45.590$, $SE = 19.840$, $t = -2.297$, $p = 0.031$; Appendix, Table B5). Similarly, the perception model showed that the position of the perceptual categories for /y/ moved to a more posterior position with smaller F2 values after conflicting orthography had been introduced (β = −109.450, *SE* = 33.830, *t* = −3.235, *p* = 0.004; Appendix, Table B6). These results, visualized in Fig. 5, show that both production and perception of /y/ became less target-like after exposure to conflicting orthographic forms.

3.2.2. Production and perception of /ɛ/

Two linear mixed-effects models were conducted. The production model had F1 production as the dependent variable and the perception model had the F1 of tokens categorized as /ɛ/ as the dependent variable. The remaining model structure was identical to the models for /y/ reported above (model R code: F1 ~ Time + (1 + Time | Participant)). In line with the results reported above for /ɛ/, neither model detected a significant change in F1 over time (Appendix, Tables B7 and B8). Fig. 6 visualizes produced and perceived F1 before and after exposure to orthographic forms.

4. Discussion

This study investigated the effects of conflicting and novel orthographic forms on non-native speech sound learning. In a multi-session vowel learning study, L1-Spanish speakers were exposed to the non-native vowels /y/ and /ɛ/. Participants were either exposed to orthographic forms that were incongruent with L1 (/y/ = <u>, /ɛ/ = <e>; Conflicting Orthography group), novel orthographic forms (/y/ = <ɸ>, /ɛ/ = <ɘ>; Novel Orthography group), or no orthographic forms (Auditory-only control group). After three exposure sessions, participants were tested on their vowel production in an elicited production

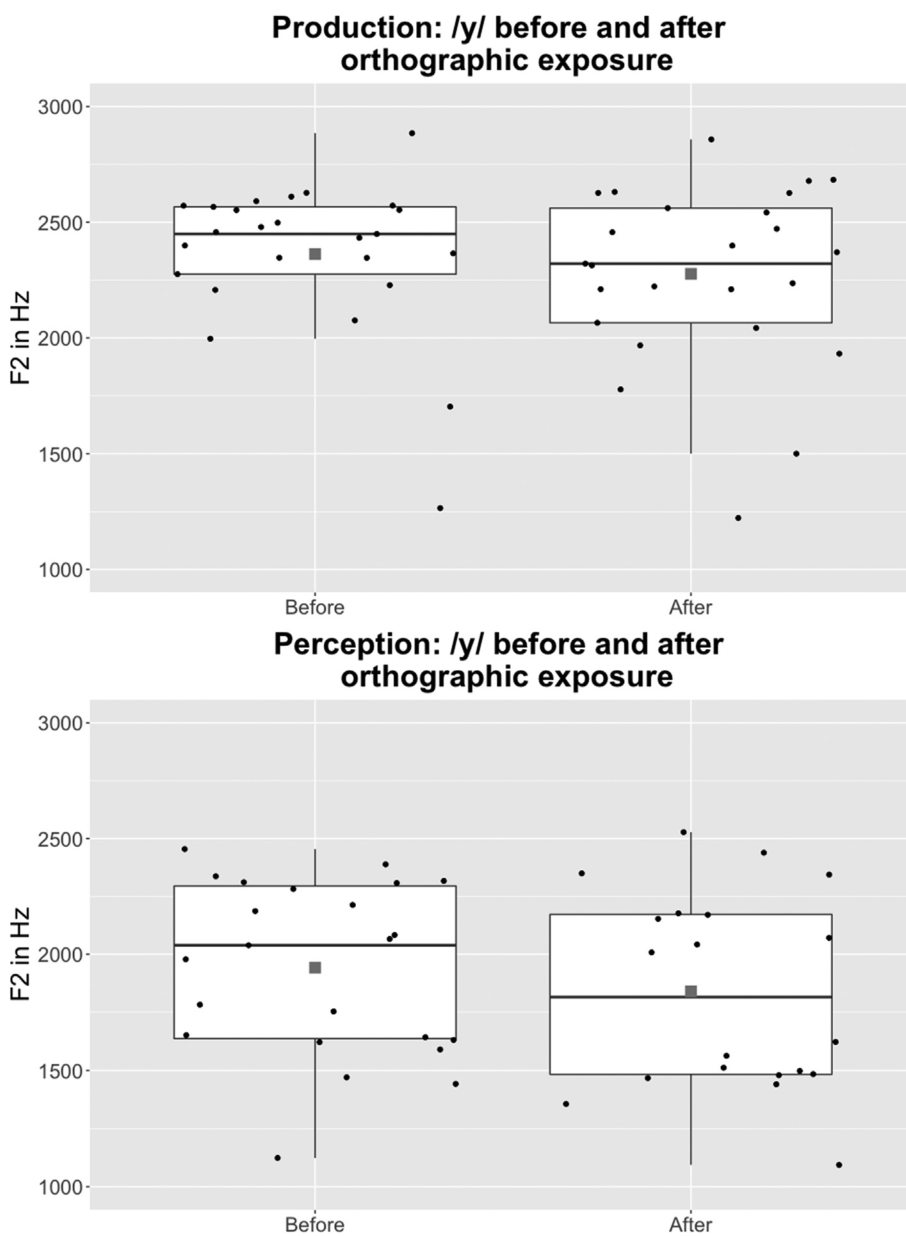

Fig. 5. /y/ production (top) and /y/ perception (bottom) before and after exposure to orthographic forms. The boxes show the first and third quartiles, the line indicates the median, the square shows the mean, and each dot represents an individual participant.

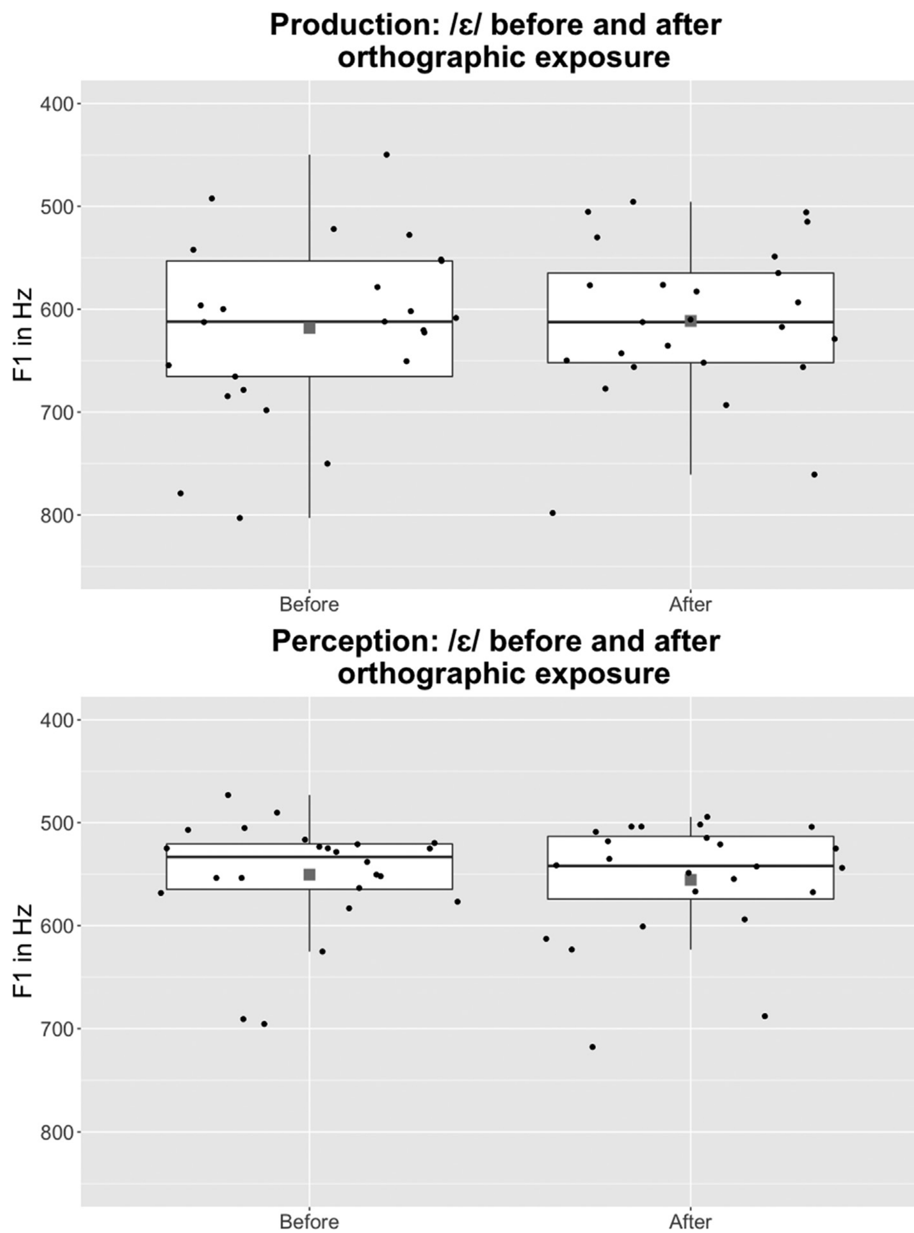

Fig. 6. /ɛ/ production (top) and /ɛ/ perception (bottom) before and after exposure to orthographic forms. The boxes show the first and third quartiles, the line indicates the median, the square shows the mean, and each dot represents an individual participant.

task and on their vowel perception in a multiple forced choice task. Participants in the Auditory-only group returned for a fourth session, in which they were exposed to conflicting orthography for /y/ and novel orthography for /ɛ/. Subsequently, they were retested on their vowel production and perception.

4.1. The effect of conflicting and novel orthographic forms on vowel production and perception

Acquisition of new categories was defined as formant differences between the newly learned vowels and their neighboring native vowels. We assessed F2 differences between /y/ versus L1 /i/ and /u/ and the F1 difference between /ɛ/ versus L1 /e/. All groups formed new categories for /y/ and /ɛ/ in production and perception, suggesting that they had perceived /y/ and /ɛ/ as *new phones* according to the SLM (Flege, 1995) or as *uncategorized sounds* according to the PAM-L2 (Best & Tyler, 2007) frameworks.

Participants' formation of an /ɛ/ category regardless of the learning condition was unexpected: Given the reported similarity of /ɛ/ to Spanish /e/ (Kartushina & Frauenfelder, 2014; Kartushina & Martin, 2019; Pallier et al., 1997; Sebastián-Gallés et al., 2005, 2006, 2009), we hypothesized that participants receiving either conflicting orthographic input or auditory-only input would map /ɛ/ onto their L1 /e/ category. This would be attributed to *equivalence classification* in the SLM (Flege, 1995) or *perceptual assimilation* in the PAM-L2 (Best & Tyler, 2007). We further hypothesized that having a link to a novel orthographic form would help participants detect small acoustic differences, in particular between /ɛ/ and L1 /e/, predicting that the Novel Orthography group would establish a new category for /ɛ/. Neither of these hypotheses were confirmed: All three groups established a new /ɛ/ category in production and perception, suggesting that the small acoustic differences between /ɛ/ and /e/ were sufficiently large for all participants to perceive /ɛ/ as a new phone (SLM; Flege, 1995) or uncategorized sound (PAM-L2; Best & Tyler, 2007). Note that the instructions and settings for the experiment may have discouraged equivalence classification/perceptual assimilation. Participants had been invited to participate in an experiment on non-native sound learning; they were aware that they would be learning non-native vowels. This may have encouraged them to pay more attention to the small acoustic differences between /ɛ/ and L1 /e/ than they might have done in a word learning task. There was also a difference between the non-native vowel perception task and the L1 vowel perception task: the non-native vowel perception task included two vowel response categories and one corresponding to 'none of these options', whereas the L1 vowel perception task had five vowel response categories (plus 'none of these options'), suggesting that the observed differences between novel vowel and L1 vowel categories may be related to task differences. Yet, the same 613 stimuli were used in both tasks, therefore, the range of acoustic features presented in both tasks was identical. In sum, although the results showed that all participants formed new categories for /ɛ/ in production and perception, this difference between response categories in the perception task must be acknowledged.

In contrast, the newly formed production and perception categories for /y/ differed in position depending on learning conditions. In particular, the presence of conflicting orthographic forms during auditory exposure to /y/ led to production and perception categories with smaller F2, indicating a more posterior position closer to /u/ than those formed with auditory-only exposure. Importantly, the Conflicting Orthography groups' productions covered the entire F2 range from 500 to 3000 Hz, with their production including tokens typical for L1 /u/, novel /y/, and those falling in between the two.

The present study extends previous research on the effect of incongruent GPCs at the word level, demonstrating that incongruent GPCs can also affect production of isolated speech sounds without any lexical context (Bassetti, 2017; Bassetti & Atkinson, 2015; Bassetti et al., 2018; Bürki et al., 2019; Cerni et al., 2019; Nimz & Khattab, 2020; Rafat, 2015; Showalter & Hayes-Harb, 2015; Welby et al., 2021; Young-Scholten & Langer, 2015). The effect of incongruent GPCs on the perception of consonants embedded in syllables without lexical context was recently found to cause slower response times in auditory discrimination but not lower accuracy in L2-Basque speakers (Stoehr & Martin, 2022). The present study demonstrates that incongruent GPCs enhance interaction between native and non-native vowels at the segmental levels of production and perception. This suggests that the presence of the grapheme <u> during exposure activated the L1 vowel /u/ and led to a more posterior category for the newly learned vowel /y/, providing evidence that orthographic information may be encoded at the speech sound level. As a consequence, orthographic effects at the lexical level may not result exclusively from co-activation of L1 orthographic forms during L2 word processing but may also be driven by orthographic influence on speech sound representations.

Nevertheless, given that this study focused on isolated vowels and did not additionally include words, we cannot assume that the present results would generalize to lexical items containing these critical vowels. Strong effects of orthographic incongruence between L1 and L2 segments in word production, however, have been reported previously (Bassetti, 2017; Bassetti & Atkinson, 2015; Bassetti et al., 2018; Bürki et al., 2019; Cerni et al., 2019; Nimz & Khattab, 2020; Rafat, 2015; Showalter & Hayes-Harb, 2015; Welby et al., 2021; Young-Scholten & Langer, 2015). Furthermore, the present study involved L1 speakers of Spanish, which has highly transparent GPC mappings. It is possible that the observed effects of incongruent GPCs on vowel production and perception are limited to speakers of such transparent languages and may not extend to speakers of orthographically opaque languages. Yet, given that L1 speakers of orthographically opaque English (e.g., Rafat, 2015) and French (e.g., Bürki et al., 2019) show orthographic effects in L2 word production when L1-L2 GPCs mismatch, we speculate that we would observe similar – but possibly weaker – effects in these populations.

Unexpectedly, in contrast to learning the new /y/ vowel category, none of the learning conditions had a detectable influence on the formation of /ɛ/ category positions. There are two possible explanations for the diverging results for /y/ and /ɛ/: differences in the learning task; and treating /ɛ/ but not /y/ as an L1 vowel. First, the different impacts of orthography on /y/ and /ɛ/ may be explained by differences in the phonetic space of these two vowels which affected the learning task demands on participants. Learning /y/ required participants to form a new category that is far away from L1 vowel categories and is located in an unoccupied area of the vowel space. Thus, participants were able to form a new /y/ category that does not overlap with L1 vowels. Learning /ɛ/ required forming a new category that is only slightly different from the L1 vowel /e/, and these two categories may even overlap (Meunier et al., 2003; Morrison, 2004). The numerically larger effect sizes and category distances found here between /y/ and both /u/ and /i/ in production and perception compared to the category distance between /ɛ/ and /e/ in production and especially perception reflect these differences in the learning task between /y/ and /ɛ/. This suggests the acoustic similarity and the lack of space between /ɛ/ and L1 /e/ may have made it more challenging for L1-Spanish speakers to learn /ɛ/ than to learn /y/ as a category distinct from L1 vowels. Research reporting that Spanish speakers have difficulty producing and perceiving the /ɛ/-/e/ contrast in French and Catalan supports the assumption that this sound contrast poses particular challenge for L1-Spanish speakers (Kartushina & Frauenfelder, 2014; Kartushina & Martin, 2019; Pallier et al., 1997; Sebastián-Gallés et al., 2005, 2006, 2009). Given the limited vowel space that /ɛ/ can occupy, it is possible that conflicting orthography played a negligible additional role in /ɛ/ production and perception. It has recently been argued that the lack of orthographic effects found in English speakers who learn the Arabic /k/-/q/ contrast may be due to the perceived closeness of these two consonants, which overrides any potential effects from orthographic input (Showalter & Hayes-Harb, 2015).

The second reason that may have caused the diverging results for /y/ and /ɛ/ is that participants considered /ɛ/ an extreme exemplar of /e/ rather than a distinct sound. If so, it would have made the orthographic mismatch ineffective. To address this possibility in future research, one could associate /y/ with the grapheme <i> instead of <u>. In Spanish, <i> corresponds to the vowel /i/, which is located adjacent to /y/ in the vowel space. If no orthographic effects were to be observed with this arrangement, it would suggest that orthographic effects may be limited to learning new vowels that are sufficiently distant from native categories and are therefore unlikely to be considered L1 exemplars. In sum, it appears that the effect of conflicting orthography on non-native sound learning is not universal, but instead depends on the nature of specific contrasts between a non-native sound and an already existing L1 category or categories.

Participants in the Novel Orthography group benefitted from the association between /y/ and a novel orthographic form in perception, as their perceptual category for /y/ was more target-like than that of the Auditory-only group. Yet, no differences between the Novel Orthography and Auditory-only groups were found either for /y/ production or for /ɛ/ production and perception. The lack of a general novel orthography advantage is in line with studies on the role of novel orthographic input in word learning, which did not detect any superiority for learning with novel orthographic input compared to auditory learning alone (Hayes-Harb & Hacking, 2015; Pytlyk, 2011; Showalter; 2018; Showalter & Hayes-Harb, 2015). However, since novel orthographic forms aided participants' perception of /y/, we speculate that over a more extended period, they may eventually aid production as well.

In sum, there is no strong evidence for a systematic benefit from the additional presence of novel orthographic input over solely auditory input on speech sound learning. Orthographic input involving incongruent L1-L2 GPCs, however, may increase interaction between native and non-native vowels, leading to less target-like production and perception. Importantly, this deviation from target vowels is not random. Instead, a non-native vowel is influenced by the L1 vowel with which it shares a grapheme, leading to positional changes in the vowel space toward the L1 vowel.

In conclusion, we consider the observed orthographic effects on mismatched orthography between L1 and L2 on the production and perception of /y/ but not /ɛ/ as a starting point for broader research in this area. In particular, a more thorough investigation of orthographic effects on speech sound learning in different populations and comprising more sounds is needed to formulate general predictions about orthographic effects on L2 speech sound learning. Given the prominence of orthography in modern societies, and its prominence in second language teaching, these predictions should be integrated into models of L2 speech production and perception.

4.2. The effect of sequential exposure to novel and conflicting orthographic forms on speech production and perception

Participants in the Auditory-only group were exclusively exposed to auditory input during the first three learning sessions. In a fourth session, the same participants were exposed to conflicting orthography for /y/, which caused a position change in their already established production and perception categories, and to novel orthography for /ɛ/, which did not affect previously formed production or perception category positions.

This lack of positional changes found in /ɛ/ production and perception categories after sequential exposure to a novel orthographic form is in line with the results of the between-group comparison after three exposure sessions, in which we found no effect of novel orthographic forms on /ɛ/ category positions in either production or perception (see Section 3.1). As the experimental design of the present study did not comprise a second sequential-exposure group with the opposite orthography distribution (i.e., novel orthography for /y/ and conflicting orthography for /ɛ/), it remains to be tested whether the absence of an effect for /ɛ/ was due to the difficulty of the vowel itself or instead reflected similar learning outcomes for simultaneous and sequential exposure to novel orthographic forms.

Sequential exposure to conflicting orthographic forms during the fourth session caused a positional shift in /y/ production and perception categories towards /u/, which further highlights the strong impact incongruent L1-L2 GPCs can have on speech sound learning: Participants were exposed to the orthographic link between /y/ and <u> for less than 40 minutes, and yet this was sufficient for their presumably already established /y/ production and perception categories to shift towards /u/. Rapid effects of incongruent GPCs on word production in a new language after short exposure phases have previously been reported (Rafat, 2015). Here, we demonstrated the effects of sequential exposure to orthographic forms in novice learners in controlled speech sound learning. More research is required to understand if this finding would be replicated in more natural learning situations. In sum, the present findings show that sequential exposure to orthographic forms incongruent with L1 affects production and perception categories that had initially been formed without orthographic input.

5. Conclusion

The present study showed that grapheme to phoneme correspondences (GPCs) may fundamentally affect the production and perception of newly learned isolated speech sounds without lexical context. While the association between novel orthographic forms and newly learned speech sounds yielded no general learning benefit in terms of phonetic precision in production and perception, incongruent GPCs between the L1 and newly learned speech sounds appeared to have detrimental effects on the phonetic characteristics of production and perception. This suggests that orthographic representations are not exclusively encoded at the lexical level but also at the speech sound level; the presence of the grapheme <u> was sufficient to activate the corresponding L1 vowel /u/. Yet, the effect of incongruent GPCs on speech sound production and perception does not appear to be universal. Speech sounds that impose a particular difficulty for learners, as has been reported for /ɛ/ in Spanish speakers, may not be further impacted by L1-L2 orthographic incongruence. Moreover, simultaneous learning of non-native speech sounds and orthographic forms and sequential exposure to orthographic forms led to similar results in production and perception. Indeed, across all measures, results were consistent for production and perception, suggesting that both modalities are equally affected by incongruent orthographic forms. In conclusion, despite the unarguable predominance of speech input, the relationship between L1 and L2 orthographic forms appears crucial for the production and perception of non-native speech sounds.

Acknowledgements

We are grateful to Itziar Basterra, Florent Dueme, and Amets Esnal for administering the present study and to Magda Altman and Caroline Handley for proofreading the manuscript. We would like to thank the members of the Speech and Bilingualism group and the Spoken Language group at the Basque Center on Cognition, Brain and Language for valuable discussions and input on this study.

This work was supported by institutional grants from the Basque Government [BERC 2018–2021 program] and the Spanish State Research Agency [BCBL Severo Ochoa excellence accreditation CEX2020-001010-S] awarded to the BCBL. This project has also received funding from the European Union's H2020 research and innovation program (Marie Skłodowska-Curie grant agreement No 843533 awarded to AS); the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 819093 to CDM); the Spanish Ministry of Economy and Competitiveness [PSI2017-82941-P; Europa-Excelencia ERC2018-092833; RED2018-102615-T (all CDM)]; and the Basque Government [PIBA18-29 (CDM)].

Appendix A. Description of the inhibition and reading tasks.

During the second learning session, participants were tested on their inhibitory and reading skills to ensure that the three groups were matched across these measures. The two tasks are described below.

Inhibition. In the retrieval induced inhibition task (Darcy et al., 2016; Lev-Ari & Peperkamp, 2013), participants had to memorize six words each from three categories (animal, occupation, and vegetable names). Three items each from two of these categories were then practiced in a recall task, and participants had to inhibit the remaining three words in each category; the six items from the excluded category served as a control. After the practice phase, participants were tested on their recognition of the 18 initially memorized words, which had been intermixed with 18 fillers. The inhibition score was calculated as the median RT for inhibited items during the recognition phase divided by the median RT for control items. A score above 1 indicated inhibition, with larger scores indicating superior inhibitory skill.

Reading. In the Alondra reading test (Lallier et al., 2022), participants had to read a meaningless text consisting of 278 words aloud. Participants' reading skills was then assessed as the number of words read correctly in a pre-determined time frame; in the current study, they were given 90 seconds.

Appendix B. Output of the linear mixed-effects models.

Table B1
/y/ production.

	β	*SE*	*t*	*p*
(Intercept)	1963.750	18.930	103.724	<0.001
Vowel (/y/ vs /i/)	−791.430	20.150	−39.280	<0.001
Vowel (/y/ vs /u/)	1096.030	27.750	39.490	<0.001
Group (Conflicting vs Auditory-only)	−70.710	26.770	−2.641	0.010
Group (Novel vs Auditory-only)	24.480	26.770	0.914	0.364
Vowel (/y/ vs /i/) * Group (Conflicting vs Auditory)	−74.130	28.490	−2.601	0.011
Vowel (/y/ vs /i/) * Group (Novel vs Auditory)	13.680	28.490	0.480	0.633
Vowel (/y/ vs /u/) * Group (Conflicting vs Auditory)	−17.750	39.250	−0.452	0.653
Vowel (/y/ vs /u/) * Group (Novel vs Auditory)	30.640	39.250	0.781	0.438

Table B2
/y/ perception.

	β	*SE*	*t*	*p*
(Intercept)	1699.560	16.630	102.200	<0.001
Vowel (/y/ vs /i/)	−695.100	20.200	−34.413	<0.001
Vowel (/y/ vs /u/)	829.600	20.000	41.487	<0.001
Group (Conflicting vs Auditory-only)	−81.830	23.380	−3.500	<0.001
Group (Novel vs Auditory-only)	61.610	23.650	2.605	0.011
Vowel (/y/ vs /i/) * Group (Conflicting vs Auditory)	−112.410	28.480	−3.948	<0.001
Vowel (/y/ vs /i/) * Group (Novel vs Auditory)	55.280	28.700	1.926	0.058
Vowel (/y/ vs /u/) * Group (Conflicting vs Auditory)	−64.070	28.130	−2.278	0.026
Vowel (/y/ vs /u/) * Group (Novel vs Auditory)	74.690	28.440	2.626	0.011

Table B3
/ɛ/ production model.

	β	*SE*	*t*	*p*
(Intercept)	570.820	6.016	94.877	<0.001
Vowel	45.213	5.091	8.881	<0.001
Group (Conflicting vs Auditory-only)	7.566	8.508	0.889	0.377
Group (Novel vs Auditory-only)	−6.892	8.508	−0.810	0.421
Vowel * Group (Conflicting vs Auditory-only)	−4.995	7.200	−0.694	0.490
Vowel * Group (Novel vs Auditory-only)	2.142	7.200	0.297	0.767

Table B4
/ɛ/ perception model.

	β	SE	t	p
(Intercept)	531.938	4.388	121.230	<0.001
Vowel	16.140	2.978	5.420	<0.001
Group (Conflicting vs Auditory-only)	−3.012	6.162	−0.489	0.626
Group (Novel vs Auditory-only)	−0.623	6.227	−0.100	0.921
Vowel * Group (Conflicting vs Auditory-only)	1.061	4.181	0.254	0.800
Vowel * Group (Novel vs Auditory-only)	0.706	4.226	0.167	0.868

Table B5
/y/ production before and after exposure to conflicting orthography.

	β	SE	t	p
(Intercept)	2325.460	67.890	34.254	<0.001
Time	−45.590	19.840	−2.297	0.031

Table B6
/y/ perception before and after exposure to conflicting orthography.

	β	SE	t	p
(Intercept)	1791.220	90.950	19.694	<0.001
Time	−109.450	33.830	−3.235	0.004

Table B7
/ɛ/ production before and after exposure to conflicting orthography.

	β	SE	t	p
(Intercept)	616.582	16.309	37.807	<0.001
Time	−6.271	6.960	−0.901	0.377

Table B8
/ɛ/ perception before and after exposure to conflicting orthography.

	β	SE	t	p
(Intercept)	552.670	10.780	51.282	<0.001
Time	2.860	3.850	0.743	0.464

References

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., & Atkinson, N. (2015). Effects of orthographic forms on pronunciation in experienced instructed second language learners. *Applied Psycholinguistics, 36*(1), 67–91. https://doi.org/10.1017/S0142716414000435.

Bassetti, B., Sokolović-Perović, 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*(4), 577–597. https://doi.org/10.1177/0023830918780141.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01.

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In M. J. Munro & O.-S. Bohn (Eds.), *Language experience in second language speech learning: In honor of James Emil Flege* (pp. 13–34). Benjamins. 10.1075/lllt.17.07bes.

Boersma, P., & Weenink, D. (2017). Praat: Doing phonetics by computer. Version 6.0.24. [computer program]. Available from http://www.praat.org/.

Brennan, E., Ryan, E., & Dawson, W. (1975). Scaling of apparent accentedness by magnitude estimation and sensory modality matching. *Journal of Psycholinguistic Research, 4*(1), 27–36. https://doi.org/10.1007/BF01066988.

Bürki, A., Welby, P., Clément, M., & Spinelli, E. (2019). Orthography and second language word learning: Moving beyond "friend or foe"? *Journal of the Acoustical Society of America, 145*(4), EL265–EL271. https://doi.org/10.1121/1.5094923.

Cerni, T., Bassetti, B., & Masterson, J. (2019). Effects of orthographic forms on the acquisition of novel spoken words in a second language. *Frontiers Communication, 4*, Article 31. https://doi.org/10.3389/fcomm.2019.00031.

Chládková, K., & Escudero, P. (2012). Comparing vowel perception and production in Spanish and Portuguese: European versus Latin American dialects. *Journal of the Acoustical Society of America, 131*(2), EL119–EL125. https://doi.org/10.1121/1.3674991.

Darcy, I., Mora, J. C., & Daidone, D. (2016). The role of inhibitory control in second language phonological processing. *Language Learning, 66*(4), 741–773. https://doi.org/10.1111/lang.12161.

Davis, M. H., Di Betta, A. M., Macdonald, M. J., & Gaskell, M. G. (2009). Learning and consolidation of novel spoken words. *Journal of Cognitive Neuroscience, 21*(4), 803–820. https://doi.org/10.1162/jocn.2009.21059.

de Anda, R. G. (2013). *Acoustic analysis of the allophones of the Spanish mid-front vowel/e* [unpublished doctoral dissertation]. The University of Texas at El Paso.

de Bruin, A., Carreiras, M., & Duñabeitia, J. A. (2017). The BEST dataset of language proficiency. *Frontiers in Psychology, 8*. https://doi.org/10.3389/fpsyg.2017.00522 522.

Dumay, N., & Gaskell, M. G. (2007). Sleep-associated changes in the mental representation of spoken words. *Psychological Science, 18*(1), 35–39. https://doi.org/10.1111/j.1467-9280.2007.01845.x.

Escudero, P., Boersma, P., Schurt Rauber, A., & Bion, R. A. H. (2009). A cross-dialect acoustic description of vowels: Brazilian and European Portuguese. *Journal of the Acoustical Society of America, 126*(3), 1379–1393. https://doi.org/10.1121/1.3180321.

Escudero, P., Simon, E., & Mulak, K. E. (2014). Learning words in a new language: Orthography doesn't always help. *Bilingualism: Language and Cognition, 17*(2), 384–395. https://doi.org/10.1017/S1366728913000436.

Escudero, P., & Wanrooij, K. (2010). The effect of L1 orthography on non-native vowel perception. *Language and Speech, 53*(3), 434–465. https://doi.org/10.1177/0023830910371447.

Ferguson, C., & Garnica, O. (1975). Theories of phonological development. In E. H. Lenneberg & E. Lenneberg (Eds.), *Foundations of language development* (pp. 153–180). Academic Press.

Flege, J. E. (1980). Phonetic approximation in second language acquisition. *Language Learning, 30*(1), 117–134. https://doi.org/10.1111/j.1467-1770.1980.tb00154.x.

Flege, J. E. (1981). The phonological basis of foreign accent: A hypothesis. *TESOL Quarterly, 15*(4), 443–445. https://doi.org/10.2307/3586485.

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., Birdsong, D., Bialystok, E., Mack, M., Sung, H., & Tsukada, K. (2006). Degree of foreign accent in English sentences produced by Korean children and adults. *Journal of Phonetics, 34*(2), 153–175. https://doi.org/10.1016/j.wocn.2005.05.001.

Hayes-Harb, R., & Barrios, S. (2021). The influence of orthography in second language phonological acquisition. *Language Teaching. Advanced online publication*. https://doi.org/10.1017/S0261444820000658.

Hayes-Harb, R., & Cheng, H.-W. (2016). The influence of the Pinyin and Zhuyin writing systems on the acquisition of Mandarin word forms by native English speakers. *Frontiers in Psychology, 7*, Article 785. 10.3389/fpsyg.2016.00785.

Hayes-Harb, R., & Hacking, J. F. (2015). The influence of written stress marks on native English speakers' acquisition of Russian lexical stress contrasts. *Slavic and East European Journal, 59*(1), 91–109. https://doi.org/10.30851/59.1.005.

Kartushina, N., & Frauenfelder, U. H. (2014). On the effects of L2 perception and of individual differences in L1 production on L2 pronunciation. *Frontiers in Psychology, 5*, Article 1246. 10.3389/fpsyg.2014.01246.

Kartushina, N., & Martin, C. D. (2019). Talker and acoustic variability in learning to produce nonnative sounds: Evidence from articulatory training. *Language Learning, 69*(1), 71–105. https://doi.org/10.1111/lang.12315.

Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman brief intelligence test KBIT-2.* Pearson.

Lallier, M., Galparsoro, N., Redondo, M., & Carreiras, M. (2022). *Spanish adaptation of the French Alouette test (Lefavrais, 1967)*. Manuscript in preparation.

Lefavrais, P. (1967). *Test de l'alouette: manuel.* Les éditions du centre de psychologie appliquée.

Lefavrais, P. (2005). *Alouette-R*. Les éditions du centre de psychologie appliquée.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software, 82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13.

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods, 44*(2), 325–343. https://doi.org/10.3758/s13428-011-0146-0.

Lev-Ari, S., & Peperkamp, S. (2013). Low inhibitory skill leads to non-native perception and production in bilinguals' native language. *Journal of Phonetics, 41*(5), 320–331. https://doi.org/10.1016/j.wocn.2013.06.002.

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.

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods, 44*(2), 314–324. https://doi.org/10.3758/s13428-011-0168-7.

Meunier, C., Frenck-Mestre, C., Lelekov-Boissard, T., & Le Besnerais, M. (2003). In *Production and perception of foreign vowels: Does the density of the system play a role?* (pp 723–726). Autonomous University of Barcelona.

Morrison, G.-S. (2004). An acoustic and statistical analysis of Spanish mid-vowel allophones. *Estudios de Fonética Experimental, 13*, 12–37.

Navarro Tomás, T. (1918). *Manual de pronunciación española*. Madrid: Publicaciones de la revista de filología española.

Navarro Tomás, T. (1965). Review of acoustic vowel loops of two Spanish idiolects. *Nueva Revista Filológica Hispánica, 14*, 342–345.

Nimz, K., & Khattab, G. (2020). On the role of orthography in L2 vowel production: The case of Polish learners of German. *Second Language Research, 36*(4), 623–652. https://doi.org/10.1177/0267658319828424.

Pallier, C., Bosch, L., & Sebastián-Gallés, N. (1997). A limit on behavioral plasticity in speech perception. *Cognition, 64*(3), B9–B17. https://doi.org/10.1016/S0010-0277(97)00030-9.

Piske, T., MacKay, I., & Flege, J. E. (2001). Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics, 29*(2), 191–215. https://doi.org/10.1006/jpho.2001.0134.

Pytlyk, C. (2011). Shared orthography: Do shared written symbols influence the perception of L2 sounds? *The Modern Language Journal, 95*(4), 541–557. https://doi.org/10.1111/j.1540-4781.2011.01244.x.

R Core Team (2022). *R: A language and environment for statistical computing (Version 4.2.0) [Computer software]*. Vienna, Austria. R-project.org/: R Foundation for Statistical Computing.

Racine, I., & Detey, S. (2019). Production of French close rounded vowels by Spanish learners: A corpus-based study. In M. Gibson & J. Gil Fernandez (Eds.), *Romance phonetics and phonology* (pp. 381–394). Oxford University Press. https://doi.org/10.1093/oso/9780198739401.003.0019.

Rafat, Y. (2015). The interaction of acoustic and orthographic input in the acquisition of Spanish assibilated/fricative rhotics. *Applied Psycholinguistics, 36*(1), 43–66. https://doi.org/10.1017/s0142716414000423.

RStudio Team (2022). *RStudio: Integrated Development Environment for R. RStudio* (Version IDE) [Computer software]. PBC. http://www.rstudio.com/.

Saito, K. (2015). The role of age of acquisition in late second language oral proficiency attainment. *Studies in Second Language Acquisition, 37*(4), 713–743. https://doi.org/10.1017/S0272263115000248.

Scovel, T. (1969). Foreign accent, language acquisition and cerebral dominance. *Language Learning, 19*(3–4), 245–254. https://doi.org/10.1111/j.1467-1770.1969.tb00466.x.

Sebastián-Gallés, N., Echeverría, S., & Bosch, L. (2005). The influence of initial exposure on lexical representation: Comparing early and simultaneous bilinguals. *Journal of Memory and Language, 52*(2), 240–255. https://doi.org/10.1016/j.jml.2004.11.001.

Sebastián-Gallés, N., Rodríguez-Fornells, A., de Diego-Balaguer, R., & Diaz, B. (2006). First-and second-language phonological representations in the mental lexicon. *Journal of Cognitive Neuroscience, 18*(8), 1277–1291. https://doi.org/10.1162/jocn.2006.18.8.1277.

Sebastián-Gallés, N., Vera-Constán, F., Larsson, J. P., Costa, A., & Deco, G. (2009). Lexical plasticity in early bilinguals does not alter phoneme categories: II. Experimental evidence. *Journal of Cognitive Neuroscience, 21*(12), 2343–2357. https://doi.org/10.1162/jocn.2008.21152.

Showalter, C. E. (2018). Impact of Cyrillic on native English speakers' phono-lexical acquisition of Russian. *Language and Speech, 61*(4), 565–576. https://doi.org/10.1177/0023830918761489.

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. https://doi.org/10.1177/0267658313480154.

Showalter, C. E., & Hayes-Harb, R. (2015). Native English speakers learning Arabic: The influence of novel orthographic information on second language phonological acquisition. *Applied Psycholinguistics, 36*(1), 23–42. https://doi.org/10.1017/S0142716414000411.

Simon, E., Chambless, D., & Alves, U. K. (2010). Understanding the role of orthography in the acquisition of a non-native vowel contrast. *Language Sciences, 32*(3), 380–394. https://doi.org/10.1016/j.langsci.2009.07.001.

Stoehr, A., Benders, T., van Hell, J. G., & Fikkert, P. (2017). Second language attainment and first language attrition: The case of VOT in immersed Dutch-German late bilinguals. *Second Language Research, 33*(4), 483–518. https://doi.org/10.1177/0267658317704261.

Stoehr, A., & Martin, C. D. (2022). Orthography affects L1 and L2 speech perception but not production in early bilinguals. *Bilingualism: Language and Cognition, 25*(1), 108–120. https://doi.org/10.1017/S1366728921000523.

Welby, P., Spinelli, E., & Bürki, A. (2021). Spelling provides a precise (but sometimes misplaced) phonological target. Orthography and acoustic variability in second language word learning. arXiv:2109.03490. 2109.03490

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer.

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. https://doi.org/10.1017/S0142716414000447.