The Effects of Congruent and Systematic Grapheme-Phoneme Correspondences on Novel Word Learning

Ka Keung Lee (ka_keung_lee@sfu.ca)
Department of Linguistics, Simon Fraser University, 8888 University Drive, Burnaby, BC V5A 1S6, Canada

H. Henny Yeung (henny_yeung@sfu.ca)
Department of Linguistics, Simon Fraser University, 8888 University Drive, Burnaby, BC V5A 1S6, Canada

Abstract

Previous studies observed a robust effect of grapheme-phoneme correspondences (GPCs) when learners are exposed to orthographic input while learning novel words. Specifically, if the novel words share the same GPC mapping with learners' native language, then this Congruency effect helps learning. Likewise, if the GPC is in a one-to-one mapping relation, this effect of Systematicity improves learning too. However, no studies have looked at the interaction of Congruency and Systematicity on word learning or explored both consonant and vowel stimuli. Here, we show no significant orthographic effect when the consonantal GPC mappings were manipulated, particularly because performance was close to the ceiling. And so, while congruent and systematic GPC mappings could facilitate learning, they were not as robust as the literature had suggested. Nevertheless, the effect of Systematicity was significant in the Vowel Sets, likely due to more exposure to the stimuli.

Keywords: grapheme-phoneme correspondence (GPC); orthography; second language acquisition

Introduction

After learning how to read and write, orthographic processing skills, such as mapping a particular letter to a specific sound, become a crucial part of how we store our languages. Orthography may, on the one hand, improve our verbal memory and phonemic awareness (e.g., Dehaene & Cohen, 2007) and assist us in organizing our auditory lexicon with additional orthographic information (c.f., Slowiaczek et al., 2003; Perre & Ziegler, 2008; Slowiaczek et al., 2014). On the other hand, it may also mislead us into memorizing the erroneous pronunciations of certain words, even though the words are from our native language (L1). For example, L1 English speakers are more likely to accept a mispronunciation of "castle" as /kæst̲l/ but not "hassle" as /hæst̲l/ because the grapheme <t>[1] presents in "castle" but absent from "hassle" (Ranbom & Connine, 2011).

In the last two decades, there has been a growing body of literature on second language acquisition investigating how orthographic input affects the learning outcome of second language (L2) words (or pseudowords). In general, the literature suggests there are four major factors contributing to whether orthographic input affects (pseudo)word acquisition: there are 1) Perceptibility, 2) Familiarity, 3) Congruency, and 4) Systematicity (Hayes-Harb & Barrios, 2021).

Perceptibility refers to whether learners can perceive the phonemes or phonemic contrasts in the new languages, and ***Familiarity*** defines whether learners are familiar with the written script used in the new languages. If learners cannot perceive the phonemic contrasts (e.g., /i/ vs. /ɪ/ for Spanish speakers, Escudero & Wanrooij, 2010; /u/ vs/ /y/ for English speakers, Simon et al., 2010), the extra orthographic input usually does not affect word acquisition. Similarly, when the written script looks too foreign (e.g., Arabic script for English speakers), learners would usually gain no benefit from the orthographic input too (e.g., Mathieu, 2016).

When phonemic contrasts become perceptible—either they exist in learners' L1 phonemic inventory or learners start realizing the distinctive features of the novel phonemes due to the improvement of L2 proficiency—and when written scripts are familiar to the learners, ***Congruency*** and ***Systematicity*** may also influence learners' acquisition. For example, if the grapheme-phoneme correspondence (GPC) mappings in a new language are congruent to learners' L1s (e.g., Cyrillic letter <K> is pronounced as /k/ for English speakers), the input tends to have a robust facilitating effect. Otherwise, if a mapping is incongruent (e.g., Cyrillic letter <H> is pronounced as /n/ instead of /h/ for English speakers), this input is detrimental (e.g., Showalter, 2018). Similarly, if the mapping is in a one-to-one corresponding systematic relation (e.g., the letter <e> is always pronounced as /ɛ/ in a pseudoword learning task; Escudero et al., 2008), then orthographic input would likely be beneficial. Inversely, if the GPC is non-systematic (e.g., German letter <a> was mapped to both /a/ and /a:/ for Polish speakers; Nimz & Khattab, 2015), the input is likely hindering.

Unlike the effects of Perceptibility and Familiarity, which are rather well-established in the literature, fewer studies, to our knowledge, have focused on the effect of Congruency or Systematicity itself without considering their interactions with Perceptibility and Familiarity, let alone considering any potential interactions between Congruency and Systematicity. The present study fills this gap in prior literature by investigating the interaction between Congruency and Systematicity while ensuring that all stimuli are both perceptible and familiar.

If Congruency and Systematicity are equally important (Possibility 1), both congruent and systematic stimuli should facilitate learning. Similarly, the inhibitory effect from

[1] In the present study, phonemes and the pronunciation of a (pseudo)word are represented in "//"; graphemes and the written form of a (pseudo)word are represented in "<>".

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society*

incongruent and non-systematic stimuli may be attenuated by the facilitating effect of Systematicity and Congruency respectively (i.e., the congruent but non-systematic stimuli will render similar learning outcomes as the incongruent but systematic ones). If Congruency is more crucial (Possibility 2), then congruent stimuli will have better learning outcomes regardless of their systematicity. Otherwise, if Systematicity is more influential for the learning outcome (Possibility 3), the systematic stimuli will lead to a better learning outcome, regardless of Congruency. Based on the robust congruent effects reported in prior literature and a facilitating systematic effect that was also observed under a congruent mapping relation (e.g., Showalter, 2018), we hypothesized that Possibility 2 (Congruency is more important than Systematicity) would more likely be the case.

Method

Participants

Participants were 282 (140 female, 135 male, 5 non-binary, 2 unknown) native English speakers from Canada (*n* = 63) and the United States (*n* = 219) between 19 and 30 years of age. The sample size was determined based on a power analysis using G*Power (Faul et al., 2007) and based on effect sizes from Barrios and Hayes-Harb (2020) with 0.95% power and an alpha of 0.05. The participants were recruited and compensated through Prolific (www.prolific.com): All were right-handed without known reading and listening disorders and completed the study on a desktop/laptop with a keyboard.

Stimuli

Between-Subjects Variables. Participants were randomly assigned to either the Audio-Only (A) group or the Orthography+Audio (OA) group. The only difference between the A and OA groups was that the participants would only receive the auditory formats of the stimuli during the learning phase in the A condition, whereas the participants would receive both the auditory and orthographic forms of the stimuli in that phase of the OA condition. Each group contained one of four sets of stimuli: Two were Consonant Sets, where the onset of the pseudowords was the target for learning, and the other two were Vowel Sets, where the nucleus was the target of learning. The stimuli used within Consonant Set 1 and 2 and within Vowel Set 1 and 2 were the same, only the combinations of their written forms (if provided) and pronunciations were different. Therefore, the study consisted of eight learning sets: 2 groups (A or OA) * 2 stimulus sets (Consonant or Vowel) * 2 sets.

Within-Subjects Variables. The target phonemes and graphemes were all perceptible (phonemes exist in English) and familiar (Roman alphabet) to the participants, as the present study only manipulated the Congruency and Systematicity of the GPC mappings. Therefore, there were 4 contexts in each set in the study (see Table 1), they were: 1) Congruent-and-Systematic (C-S); 2) Congruent-but-Non-Systematic (C-NS); 3) Incongruent-but-Systematic (IC-S); and 4) Incongruent-and-Non-Systematic (IC-NS).

Table 1: Stimuli Example.[2]

	C-S	C-NS	IC-S	IC-NS
C Set 1	<sh>-/ʃ/	<wh>-/h/ & /w/	(ch-/f/)	<ph>-/tʃ/ & /k/
	<shaas>-/ʃɑs/ <shoos>-/ʃus/	<whaas>-/wɑs/ <whoos>-/hus/	<chaas>-/fɑs/ <choos>-/fus/	<phaas>-/kɑs/ <phoos>-/tʃus/
C Set 2	<ph>-/f/	(ch-/tʃ/ & /k/)	<wh>-/ʃ/	<sh>-/h/ & /w/
	<phaas>-/fɑs/ <phoos>-/fus/	<chaas>-/kɑs/ <choos>-/tʃus/	<whaas>-/ʃɑs/ <whoos>-/ʃus/	<shaas>-/wɑs/ <shoos>-/hus/
V Set 1	<oa>-/oʊ/	<ie>-/i/ & /ɪ/)	<oo>-/eɪ/	<ea>-/u/ & /ʌ/
	<ghoas>-/goʊs/ <thoas>-/θoʊs/	<ghies>-/gɪs/ <this>-/θis/	<ghoos>-/geɪs/ <thoos>-/θeɪs/	<gheas>-/gʌs/ <theas>-/θus/
V Set 2	<ea>-/eɪ/	<oo>-/u/ & /ʌ/	<ie>-/oʊ/	<oa>-/i/ & /ɪ/
	<gheas>-/geɪs/ <theas>-/θeɪs/	<ghoos>-/gʌs/ <thoos>-/θus/	<ghies>-/goʊs/ <this>-/θoʊs/	<ghoas>-/gɪs/ <thoas>-/θis/

Consider C Set 1 as an example. Pesudowords where the grapheme <sh> was mapped to phoneme /ʃ/ used a C-S GPC mapping because this shares the same GPC mapping with English (e.g., <sh> in "shove" /ʃʌv/), and both the grapheme and phoneme are in a one-grapheme-to-one-phoneme mapping within the set. The grapheme <wh> maps to phonemes /h/ and /w/ is a set of C-NS GPC mappings because this mapping shares the same mappings with English (e.g., <wh> in "who" /hu:/ and "why" /waɪ/), but they are not in a one-to-one relation. Additionally, the grapheme <ch> to phoneme /f/ mapping is an IC-S GPC mapping because <ch> is never pronounced as /f/ in English, whereas the mapping is still one-to-one. Finally, the mappings among the grapheme <ph> and the phonemes /tʃ/ and /k/ are considered IC-NS GPC mappings because they neither share mappings with English nor are they in a one-to-one matching relation.

Critically, the stimuli used in Sets 1 and 2 were the same, only the mappings between the target graphemes and their matching phonemes were re-combined. For instance, the C-S <sh> to /ʃ/ GPC mapping in Consonant Set 1 became IC-NS <sh> to /h/ and /w/ GPC mappings in Consonant Set 2.

All stimuli in the present study were recorded in a sound-attenuated booth (at a sampling rate of 44.1 kHz) by a female native English-speaking phonetician from Canada. All stimuli were delivered using jsPsych (de Leeuw et al., 2023) through Prolific.

Procedure

The online experiment contained three parts: 1) two learning phases, 2) a criterion test, and 3) a final test. The whole experiment lasted around 15 minutes. Before the experiment, the participants received a digital copy of the consent form that could be read and downloaded to their own devices. If they consented to take part in the study, they need to press

[2] C Set = Consonant Set, V Set = Vowel Set. The orthographic forms in "<>" are the target graphemes or pseudowords used in the study, and the pronunciations in "//" are the target phonemes or auditory forms of the target graphemes or pseudowords. The underlined auditory forms have higher phonotactic probabilities.

"Continue" and then fill out a demographic survey. After that, the participants would be randomly assigned to one of the eight learning sets.

The Learning Phase. There were two blocks in the learning phase section, and the two blocks were identical except that the ordering of pseudoword stimuli was distinct, because the presentation of the stimuli was randomized. Before the learning phase, the participants received written instructions telling them they would have two tests following the learning phase, and they would have to score 90% accuracy to proceed to the final test, otherwise, the experiment would end, and they would receive half of the compensation (£1.6).

During the learning phase, the participants assigned to the A group would be exposed to the meaning (pseudo-object images) and auditory input of the target pseudowords, whereas the participants in the OA group received extra orthographic input. The meaning and auditory input (and orthographic input in the OA group) were displayed on-screen simultaneously for 3,000 ms, followed by a fixation mark "+" lasting for 500 ms, and then another target pseudoword (see Figure 1). Each pseudoword was played 4 times in each learning phase, therefore, the participants learned 64 trials across 2 phases (4 * 8 pseudowords * 2). A self-paced break was provided between the two phases.

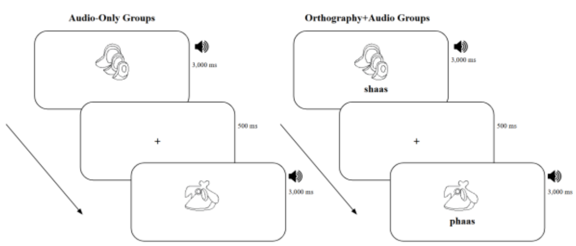

Figure 1: The Learning Phase.

The Criterion Test. Once the learning was finished, the participants were directed to the written instruction pages of the criterion test. In the criterion test, the A and AO participants only saw the meaning of each target pseudoword 4 times, together with the auditory input that either matched or mismatched the pronunciation of the learned pseudoword. The purpose of the criterion test was to ensure that the participants had learned (or reached some high threshold for achieving a basic learning effect) and paid sufficient attention during learning to proceed to the final test.

Since the criterion foils sounded very different from the target pronunciations (i.e., with differences on both onset and nucleus segments, e.g., target /ʃɑs/ vs. criterion foil /θis/), the criterion test should be relatively easy for the participants. If they failed, participants were deemed not to have paid sufficient attention during the learning phase.

The participants saw the meanings and heard either the matched or mismatched (foil) auditory input simultaneously, and they had to judge whether the auditory input and the meaning (the pseudo-object images) matched by pressing the "Z" key or not matched by pressing the "M" key on their keyboards (see Figure 2). They had 6,000 ms to respond before the page disappeared. After the participants gave their answers or the time-out, a fixation mark would appear for 500 ms followed by another judgment trial. The participants completed 32 trials (8 pseudowords * 2 matches * 2 mismatches), and they had to reach at least 90% accuracy (29 out of 32 trials) to proceed to the final test. The data of the participants who failed to pass the criterion test were excluded from the analysis.

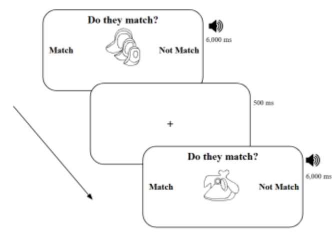

Figure 2: The Criterion/Final Test.

The Final Test. If the participants passed the criterion test, they would be directed to the final test instruction pages. The procedure for the final test was identical to that for the criterion test, but a different set of mismatched foils was used. Unlike the criterion test, the foils in the final test were different from the targets only on the target segments (onset consonants in the Consonant Sets and nucleus vowels in the Vowel Sets). Therefore, the final test was presumably more difficult.

The final test foils in the Consonant Sets differed from the targets in either place of articulation (e.g., for target /ʃɑs/ vs. foil /sɑs/, they were both voiceless and fricative consonants, but /ʃ/ is postalveolar whereas /s/ is alveolar) or manner of articulation (e.g., for target /tʃus/ vs. foil /tus/, they were both voiceless, but /tʃ/ was an affricative whereas /t/ was a plosive[3]).

Since English does not have a large phonemic inventory of pure vowels, the present study included both monophthong and diphthong vowel stimuli. The foils in the final test differed from their corresponding targets in terms of tenseness (e.g., for target /θus/ vs. foil /θʊs/, /u/ and /ʊ/ were both high back vowels, whereas /u/ was tense and /ʊ/ was lax), and height (e.g., for target /θis/ vs. foil /θæs/, both /i/ and /æ/ were both front vowels, but /i/ was high whereas /æ/ was lower). When the differences between targets and foils were between two diphthongs, the two diphthongs shared the ending vowels (e.g., for target /goʊs/ vs. foil /gaʊs/, both /oʊ/ and /aʊ/ shared the ending vowel /ʊ/). When the targets were

[3] Although /ʃ/ might be a better candidate as a foil for the target /tʃus/ because they are both postalveolar, while /t/ is alveolar. Since /ʃus/ has been used as a target and we did not want to confuse the participants and use /ʃus/ as both target and foil in the same experiment, the only available phoneme candidate left was /t/.

monophthongs and the foils were diphthongs, the ending vowel of the diphthong was the target vowel (e.g., for target /gɪs/ vs. foil /gɔɪs/, the ending vowel of the foil /ɔɪ/ was the target vowel /ɪ/). The experiment ended after the final test, and the participants who completed it were compensated with £3.2. The present study analyzed only the response data from the final test.

Results

Trials that timed out at 6,000 ms were excluded from the analysis. Since we programmed the experiment in a way that participants had to finish hearing the whole auditory input before they could respond, no trial in the present study had a reaction time (RT) that was below 200 ms, hence, no removed data point was under this condition (trials with RTs below 200). This trimming resulted in a loss of 0.17% (16 data points) of the data. All stimuli and data are available at https://osf.io/3gexs.

Analysis

To test for the presence of an orthographic effect with the interaction of Congruency and Systematicity, we ran maximally specified mixed-effects models (Baayen et al., 2018) implemented in the afex package (Singmann et al., 2024) in R (R core team, 2024). Post-hoc was run using the emmeans package (Lenth, 2025) with Bonferroni corrections.

The responses (Accuracy, with correct responses coded as 1 and incorrect responses coded as 0) were regressed in a logistic model on Orthography (whether the data was from the A or OA groups), Congruency (whether the stimuli were congruent or incongruent), Systematicity (whether the stimuli were systematic or non-systematic), and their interaction. In the original model, the random effect structures included parameters of intercept, Congruency, Systematicity, and their interaction for each participant (SubjID), and parameters of intercept, and Orthography of each word (Item, e.g., regardless of whether orthography was in C-S context or IC-NS context, it belonged to the same Item). The full model was "*Accuracy ~ Orthography * Congruency * Systematicity + (1 + Congruency * Systematicity | SubjID) + (1 + Orthography | Item)*". If the model did not converge, the random slopes and intercepts were decorrelated, and then the random slopes were moved if the model was still not converged. The simplification was repeated until the model converged.

Consonant Sets

The final converged model for Consonant Sets was "*Accuracy ~ Orthography * Congruency * Systematicity + (1 | SubjID)*". The original model had to be simplified before it converged, which might be due to the ceiling effect observed in participants' responses (see Figure 3).

Orthographic Effect. Participants assigned to the A groups performed better than those assigned to OA groups, but it was not significant (p = .19). In Figure 3, the wider the violin plot indicates more data points concentrate at the respective accuracy scores. Therefore, the figure suggests that the participants in both groups performed equally well and most of them reached the ceiling.

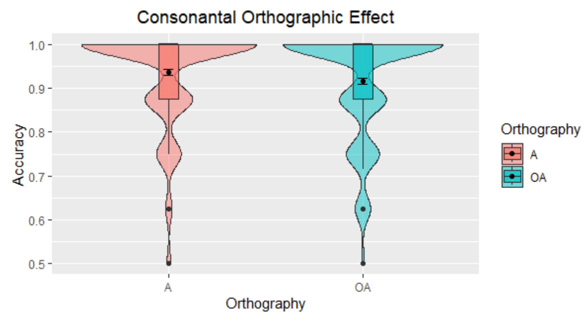

Figure 3: Consonantal Orthography Effect.[4]

Effects of Congruency and Systematicity. Congruent stimuli rendered a marginally significant facilitating effect in learning (p = .01). Systematic stimuli resulted in a numerically better learning outcome, but the effect was not significant (p = .53).

The interaction between Congruency and Systematicity was significant (p = .001). When the stimuli were non-systematic, the participants with the C-NS stimuli performed significantly better (see Figure 4) than those with IC-NS (p < .001). The effect was not observed when the stimuli were systematic (p = .27). The three-way interaction among Orthography, Congruency, and Systematicity was not significant (p = .68).

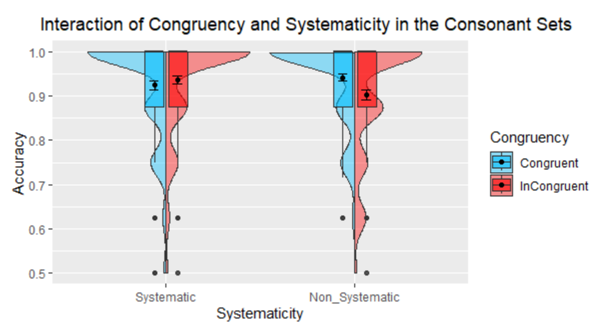

Figure 4: Interaction of Congruency and Systematicity in the Consonant Sets.[5]

In an exploratory analysis, however, the learning effect for C-NS stimuli was significantly better than IC-NS (p = .01),

[4] The salmon and cyan plots represent the results of the A and OA groups. The bar at the center of each plot is an error bar.

[5] The left and right plots represent the results of systematic and non-systematic stimuli respectively, and the split blue and split red plots represent the results of congruent and incongruent stimuli respectively. The bar at the center of each plot is an error bar.

and the one for IC-S was significantly better than IC-NS (p = .01), but these patterns were only observed in the participants who were assigned to the OA group. The participants in the A group performed equally well with either type of stimulus. Therefore, the results may suggest that when orthography was provided, the potential facilitating effect of congruent and systematic GPCs could attenuate the negative effects of incongruency and non-systematicity. However, without an omnibus interaction, interpretation of these results is limited.

Vowel Sets

The final converged model for Vowel Sets was "(*Accuracy ~ Orthography * Congruency * Systematicity + (1+ Congruency + Systematicity || SubjID) + (1+ Orthography || Item*)".

Orthographic Effects. Participants assigned to the A groups performed significantly better (p = .006) than those assigned to the OA groups (see Figure 5).

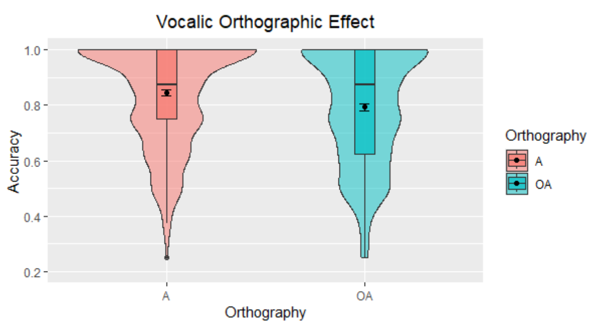

Figure 5: Vocalic Orthography Effect.[6]

Effects of Congruency and Systematicity. Unlike in Consonant Sets, congruent stimuli in Vowel Sets did not have a significant effect (p = .4), but Systematicity played a significant role in learning, where non-systematic stimuli significantly hindered ($p < .001$) the acquisition (see Figure 6). However, the two-way interaction between Systematicity and Orthography was not significant (p = .42).

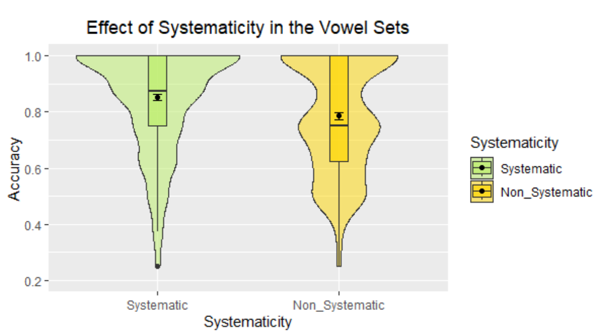

Figure 6: Effect of Systematicity in the Vowel Sets.[7]

Discussion

The present study suggests that orthographic input and the factors of the Congruency and Systematicity of GPCs between a learner's L1 and L2 may play different roles depending on whether the new GPC mappings are applied to consonant or vowel segments. In Consonant Sets, where the consonantal GPC mappings were manipulated, orthographic input did not affect the learning, and participants in both A and OA groups learned the pseudowords equally well. When the vocalic GPC mappings were manipulated in Vowel Sets, however, the participants who received orthographic input performed significantly worse than those who were only exposed to auditory stimuli. In general, we explain these effects by suggesting that GPC mappings in Consonant Sets were easier to learn compared to the GPC mappings in Vowel Sets. The ceiling effect in Consonant Sets was also reflected during the model simplification; that is to say, a random slope had to be removed to make the model converge. On the other hand, the model for Vowel Sets converged after the random slopes and intercepts were decorrelated. Overall, these results may be consistent with previous studies that argued vowels are harder to learn compared to consonants (e.g., Vergara-Martínez et al., 2011; Bouchon et al., 2015).

The finding of lack of orthographic effect in Consonant Sets is consistent with previous studies, which found a null effect with Imperceptible stimuli (e.g., Escudero & Wanrooij, 2010; Simon et al., 2010; Escudero et al., 2014) but inconsistent with studies that found either facilitating (e.g., Escudero & Wanrooij, 2010; Escudero et al., 2014) or hindering effect (e.g., Escudero & Wanrooij, 2010; Escudero et al., 2014; Stoehr & Martin, 2022) with stimuli that their participants were able to perceive (although with difficulty). The stimuli of the present study should not cause any difficulty in perception for the recruited participants, because all our phonemes exist in the English phonetic inventory. Therefore, together with previous findings, the present study provides evidence that orthographic input may play a role in learning only when the new phonemes are perceptible, but also only when there is a certain degree of perceptual difficulty (at least on consonant segments) for the learners.

Another difference that this study had with previous studies is that prior work had allowed their participants to re-take the criterion test multiple times until they reached 90% accuracy (e.g., Hayes-Harb & Cheng, 2016; Mathieu, 2016; Showalter, 2018; Barrios & Hayes-Harb, 2020), whereas the present study only allowed a single attempt. If the participants passed, they could proceed to the final test; otherwise, the experiment would end. Therefore, all participants in our final data set had only gone through the criterion test once, whereas the participants in previous studies might have gone through it several times. In this case, the potential uneven

[6] The salmon plot represents the results of the A groups, and the cyan plot represents the results of the OA groups. The bar at the center of each plot is an error bar.

[7] The green and yellow plots represent the results of systematic and non-systematic stimuli respectively. The bar at the center of each plot is an error bar.

times of exposure to the stimuli (i.e., the participants who have taken the criterion test four times had more exposure to the stimuli than those who passed the test on their first try) and individual differences (i.e., the participants who passed the test on their first try might be better at learning new words than those who had to repeat the test several times) might influence results, and consistency between prior work and the present study.

The results of Consonant Sets also revealed that congruent stimuli led to a better learning effect, but the participants did not actually perform better with congruent stimuli (possibly they were at ceiling in their correct responses); it was because they performed worse when GPCs were incongruent, especially when the stimuli were also non-systematic. However, the results might be hard to explain without taking Orthography into consideration (i.e., the factor of Congruency depends on the presence of orthography because it concerns the mapping of graphemes and phonemes). Thus, although there was not an overall orthographic effect, the exploratory analysis found that only in the condition where orthographic input was provided, the congruent GPCs could attenuate the negative effect of non-systematicity and render better learning outcomes than the incongruent GPC mappings. If the stimuli were incongruent, on the other hand, the systematic GPC mappings could also mitigate the negative incongruent effect and render better learning outcomes than non-systematic GPC mappings. In other words, when incongruent or non-systematic GPC mappings render a hindering effect, the systematic or congruent GPC mappings may attenuate their negative effect, respectively. In this case, the exploratory analysis result from Consonant Sets may be against our hypothesis and suggest that Congruency and Systematicity are equally important, at least for learning consonantal GPC mappings.

A robust effect of Congruency (i.e., facilitating effect that participants with congruent GPC input had relative to those who only received auditory input) was not replicated in the Consonant Sets analysis may also have been due to the level of difficulty of foils used in the previous studies. In the present study, the foils were similar to the target pronunciations in many ways but would not bias the participants to choose them due to Congruency. For instance, the foil for the incongruent stimulus <phaas> as /kɑs/ was /pɑs/, in which /k/ differs from /p/ in place, instead of /fɑs/, which the foil would then become congruent (i.e., <ph> as /f/). But in previous studies (e.g., Showalter, 2018), the foil of an incongruent GPC mapping (<HOM> as /nom/) was actually the congruent word (e.g., /hom/). In this case, their foils might influence their English-speaking participants to be more likely to choose the congruent foil (/hom/) instead of the incongruent target (/nom/) for their incongruent stimuli (<HOM>). In other words, the congruent effect might be real, but it is not as robust as previous studies have suggested. The strong congruent effect observed in prior literature might be a consequence of using potentially biasing foils.

The orthographic effect was replicated in Vowel Sets, and our results may suggest that learning vowels is hard, and seeing the orthography during the experiment might make the vowel learning even harder because the participants have to process the extra visual information during learning. In addition, the effect of Systematicity was observed, such that the systematic stimuli rendered better learning outcomes than the non-systematic ones. Broadly, the findings in Vowel Sets seem to be consistent with studies that observed the hindering effect of non-systematic stimuli for vowel learning (e.g., Nimz & Khattab, 2015).

However, these conclusions are limited by the lack of a significant interaction between Systematicity and Orthography. Thus, this result might not suggest an effect of systematic GPC mappings but instead reveal an important role of the amount of exposure to the target vowel phonemes. In the present study, the participants had a chance to hear the systematic phonemes (e.g., /oʊ/ in <ghoas> as /goʊs/ and <thoas> as /θoʊs/ pair) 16 times in the learning phases (2 pseudowords * 4 times each phase * 2 phases), but only 8 times for the non-systematic ones (e.g., /ɪ/ and /i/ in <ghies> as /gɪs/ and <thies> as /θis/ pair; 1 pseudoword * 4 times each phase * 2 phases). Since vowels might potentially be harder to learn, the participants might have a worse learning outcome with less exposure to the (non-systematic) vowels.

In spite of having a large sample size, a limitation of the present study is that our stimuli might be relatively easy to learn, therefore, most of the participants reached the ceiling, especially those assigned to Consonant Sets. That might be the reason that the effects and the interaction between Congruency and Systematicity were not found significantly in the present study. A follow-up study will be conducted to investigate if an orthographic effect could be observed when learning is harder, and whether the effects of Congruency and Systematicity have different roles to play when the GPC mappings are manipulated on consonants versus vowels.

To summarize, the present study suggests that the effect of Congruency and Systematicity may play a role in learning, but it may only be the case when the learners are at a disadvantage of being exposed to non-systematic and incongruent GPCs respectively. Learning vowels themselves might be inherently harder than learning consonants; therefore, spending extra cognitive resources to memorize the written forms, although the vowels are perceptible and the script is familiar, may make learning even harder. Consequently, the learners may need more exposure to achieve a decent learning outcome. Therefore, the present results suggest that exposing learners to the written forms during early stages of language acquisition might not be beneficial, but instead hinder their word acquisition.

References

Barrios, S. L., & Hayes-Harb, R. (2020). Second language learning of phonological alternations with and without orthographic input: Evidence from the acquisition of a German-like voicing alternation. *Applied Psycholinguistics, 41*(3), 517-545.

Bouchon, C., Floccia, C., Fux, T., Adda-Decker, M., & Nazzi, T. (2015). Call me Alix, not Elix: Vowels are more important than consonants in own-name recognition at 5 months. *Developmental Science, 18*(4), 587-598.

Bristow, D., Dehaene-Lambertz, G., Mattout, J., Soares, C., Gliga, T., Baillet, S., & Mangin, J. F. (2008). Hearing faces: how the infant brain matches the face it sees with the speech it hears. *Journal of cognitive neuroscience, 21*(5), 905-921.

de Leeuw, J.R., Gilbert, R.A., & Luchterhandt, B. (2023). jsPsych: Enabling an open-source collaborative ecosystem of behavioral experiments. *Journal of Open Source Software*, *8*(85), 5351, https://joss.theoj.org/papers/10.21105/joss.05351.

Dehaene, S., & Cohen, L. (2007). Cultural recycling of cortical maps. *Neuron, 56*(2), 384-398.

Escudero, P., & Wanrooij, K. (2010). The effect of L1 orthography on non-native vowel perception. *Language and speech, 53*(3), 343-365.

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., 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.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods*, *39*(2), 175-191.

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, 785.

Hayes-Harb, R., & Barrios, S. (2021). The influence of orthography in second language phonological acquisition. *Language Teaching, 54*(3), 297-326.

Lenth, R. (2025). *emmeans: Estimated Marginal Means, aka Least-Squares Means. R package* (Version 1.10.6-090003). https://rvlenth.github.io/emmeans/.

Mathieu, L. (2016). The influence of foreign scripts on the acquisition of a second language phonological contrast. *Second Language Research, 32*(2), 145-170.

Nimz, K., & Khattab, G. (2015). L2 sound perception: Does orthography matter? *International Congress of Phonetic Sciences*.

Singmann, H., Bolker, B., Westfall, J., & Aust, F. (2024). *Afex: analysis of factorial experiments* (Version 1.4-1). http://cran.r-project.org/package=afex.

Perre, L., & Ziegler, J. C. (2008). On-line activation of orthography in spoken word recognition. *Brain research*, 1188, 132-138.

Prolific. (2014). *Prolific* (Version 10.2024). Retrieved October 20, 2024, from https://www.prolific.com/.

R core team. (2024). *R: A language and environment for statistical computing* (Version 4.4.1). https://www.R-project.org/.

Ranbom, L. J., & Connine, C. M. (2011). Silent letters are activated in spoken word recognition. *Language and Cognitive Processes, 26*(2), 236-261.

Vergara-Martínez, M., Perea, M., Marín, A., & Carreiras, M. (2011). The processing of consonants and vowels during letter identity and letter position assignment in visual-word recognition: An ERP study. *Brain and language, 118*(3), 105-117.

Vitevitch, M. S., & Luce, P. A. (2004). A web-based interface to calculate phonotactic probability for words and nonwords in English. *Behavior Research Methods, Instruments, & Computers, 36*(3), 481-487.

Slowiaczek, L. M., & Kahan, T. A. (2014). Evidence for the automatic processing of prelexical codes in an orthographic but not a phonological task. *Psychonomic bulletin & review*, 21, 1607-1616.

Slowiaczek, L. M., Soltano, E. G., Wieting, S. J., & Bishop, K. L. (2003). An investigation of phonology and orthography in spoken-word recognition. *The Quarterly Journal of Experimental Psychology Section A, 56*(2), 233-262.

Showalter, C. E. (2018). Impact of Cyrillic on native English speakers' phono-lexical acquisition of Russian. *Language and Speech, 61*(4), 565-576.

Stoehr, A., & Martin, C. D. (2022). The impact of orthographic forms on speech production and perception: An artificial vowel-learning study. *Journal of Phonetics*, 94, 101180.